

Univerzita Karlova  
Přírodovědecká fakulta  
Katedra biochemie

# Studie jednovláknových DNA biofyzikálními metodami

---

## Study of single-stranded DNA by biophysical techniques

Jakub Svoboda

Vedoucí práce: doc. Ing. Bohdan Schneider, CSc.

Praha 2018





**Prohlášení**

Prohlašuji, že jsem tuto bakalářskou práci vypracoval samostatně pod vedením školitele doc. Ing. Bohdana Schneidera, CSc. a všechny použité prameny jsem řádně citoval.

V Praze dne:

Podpis:

## **Poděkování**

Rád bych poděkoval svému školiteli doc. Ing. Bohdanu Schneiderovi, CSc. za odborné vedení, cenné rady a vytvoření vynikajících pracovních podmínek na Biotechnologickém ústavu AV ČR, v. v. i.

Dále bych chtěl poděkovat Tatsianě Charnavets, Ph. D., Ing. Lucii Kolářové a Mgr. Jiřímu Zahradníkovi za pomoc a podněty při provádění experimentů a Ing. Petru Kolenkovi, Ph. D. za pomoc s měřením a řešením krystalových struktur.

## Abstrakt

DNA je základní funkční molekula všech domén života. Jednou z jejích charakteristických vlastností je asociace do tvorby dvouvláknové šroubovice. Přesto může i jednovláknová DNA tvořit nekanonické struktury jako například vlásenky, triplexy nebo tetraplexy. Mezi zajímavé jednovláknové sekvence patří REP elementy. Tyto úseky DNA jsou zástupcem bakteriálních transpozonů tzv. inserčních sekvencí, které se vyskytují ve velkém počtu kopií v široké škále bakteriálních organismů.

Tato práce se zabývá studiem struktury a charakterizací vybraných REP sekvencí. Jsou předkládány výsledky měření oligonukleotidů se sekvencemi REP elementů spektrální metodou cirkulárního dichroismu, kalorimetrickým měřením a strukturní analýzy jejich monokrystalů difrakcí paprsků X.

## Klíčová slova

Jednovláknová DNA, krystalografie, cirkulární dichroismus, kalorimetrie, transpozony, REP elementy, inserční sekvence.

## Abstract

DNA is the fundamental functional molecule of all domains of life. One of its characteristics is the ability to self-associate to form a double stranded helix. Nevertheless, even single stranded DNA can form non-canonical structures such as hairpins, triplexes or tetraplexes. One of the interesting single stranded sequences are REP elements. These sequences form a part of bacterial transposable elements, so-called insertion sequences, which can be found in a great number of copies in wide range of bacterial species.

This thesis studies structure and properties of selected REP sequences. It presents the results of spectral measurements by technique of circular dichroism, melting curves from differential scanning calorimetry, and monocrystal diffraction data.

## Key words

Single-stranded DNA, crystallography, circular dichroism, calorimetry transpozome, REP elements, insertion sequences.

# Obsah

1. Seznam zkratk .....	9
2. Teoretický úvod .....	10
2.1. Historie .....	10
2.2. Nukleotid .....	10
2.2.1.1. Báze .....	11
2.2.1.2. Ribóza.....	12
2.2.2. Glykosidická vazba .....	12
2.2.3. Cukr-fosfátová páteř .....	13
2.3. Párování a stabilita bází .....	13
2.4. Sekvence a tvorba duplexu DNA .....	14
2.4.1. Duplex jako základní strukturní a funkční forma DNA.....	15
2.5. Strukturní formy duplexu .....	16
2.6. Další DNA struktury.....	18
2.6.1. DNA trojitá šroubovice .....	19
2.6.2. Guaninový tetraplex .....	19
2.6.3. I-Motiv .....	21
2.7. Metody studia struktury nukleových kyselin.....	21
2.7.1. Difrakce X-paprsků na monokrystalech .....	21
2.7.1.1. Příprava materiálu .....	21
2.7.1.2. Příprava krystalů .....	22
2.7.1.3. Měření difrakce .....	23
2.7.1.4. Výpočet mapy elektronových hustot.....	25
2.7.1.5. Rafinace struktury .....	25
2.7.1.6. Výhody a nevýhody rentgenové strukturní analýzy .....	26
2.7.2. Spektrální metody .....	26
2.7.2.1. Nukleární magnetická rezonance.....	26
2.7.2.2. Výhody a nevýhody nukleární magnetické rezonance .....	27
2.7.2.3. Chiroptické metody .....	27
2.8. Strukturní databáze .....	27
2.9. Mobilní genetické elementy .....	28
2.9.1. Transpozony.....	28
2.9.1.1. Retrotranspozony .....	29
2.9.1.2. DNA transpozony.....	30

2.9.1.3.	REP – Repetitivní extragenní palindromické sekvence .....	31
2.9.2.	Biologický význam mobilních elementů .....	32
3.	Cíle práce .....	34
4.	Materiál a metody .....	35
4.1.	Přístroje.....	35
4.2.	Chemikálie.....	35
4.3.	Postup při přípravě krystalů oligonukleotidů .....	37
4.3.1.	Příprava oligonukleotidů na krystalizaci.....	37
4.3.2.	Tvorba krystalů s použitím krystalizačního robota a hotelu.....	37
4.3.3.	Optimalizace krystalů a fázování .....	37
4.3.4.	Měření krystalů .....	37
4.4.	Postup při měření spekter cirkulárního dichroismu.....	38
4.5.	Postup při kalorimetrickém měření .....	38
5.	Výsledky .....	39
5.1.	Analýza krystalů .....	39
5.2.	Spektra cirkulárního dichroismu .....	40
5.3.	Kalorimetrická měření .....	41
6.	Diskuze a závěr .....	44
7.	Zdroje informací .....	45



# 1. Seznam zkratek

bp – pár bazí, z anglického base pair

CD – cirkulární dichroismus

EDTA – ethylendiamintetraoctová kyselina

IS – inserční sekvence

RAYT – REP asociovaná tyrosinová transponáza

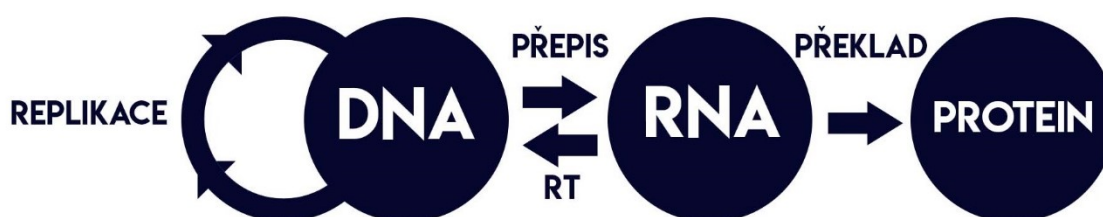
REP – repetitivní extragenní palindromická sekvence

TE – Transponovatelný element, z angl. transposable element

## 2. Teoretický úvod

### 2.1. Historie

Nukleové kyseliny jsou významné biologické polymery podílející se na uchování genetické informace (replikace), jejíž přepis (transkripce) do RNA podmiňuje překlad (translaci) genetické informace do proteinů (Obrázek 1). Tento tok genetické informace byl označen v roce 1958 jako centrální dogma molekulární biologie [1]. Struktura dvoušroubovicové DNA (dále jen duplex) byla na základě difrakčních a biochemických dat postulována v roce 1953 Jamesem Watsonem a Francisem Crickem za výrazného přispění Rosalindy Franklinové, Maurice Wilkinse, Erwina Chargaffa a jiných [2,3]. V roce 1962 byla udělena Nobelova cena za fyziologii a medicínu právě trojici Watson, Crick a Wilkins [4]



Obrázek 1: Centrální dogma molekulární biologie. Naznačen je tok genetické informace a její zachování. RT je reverzní transkripce, převzato a upraveno z [1].

### 2.2. Nukleotid

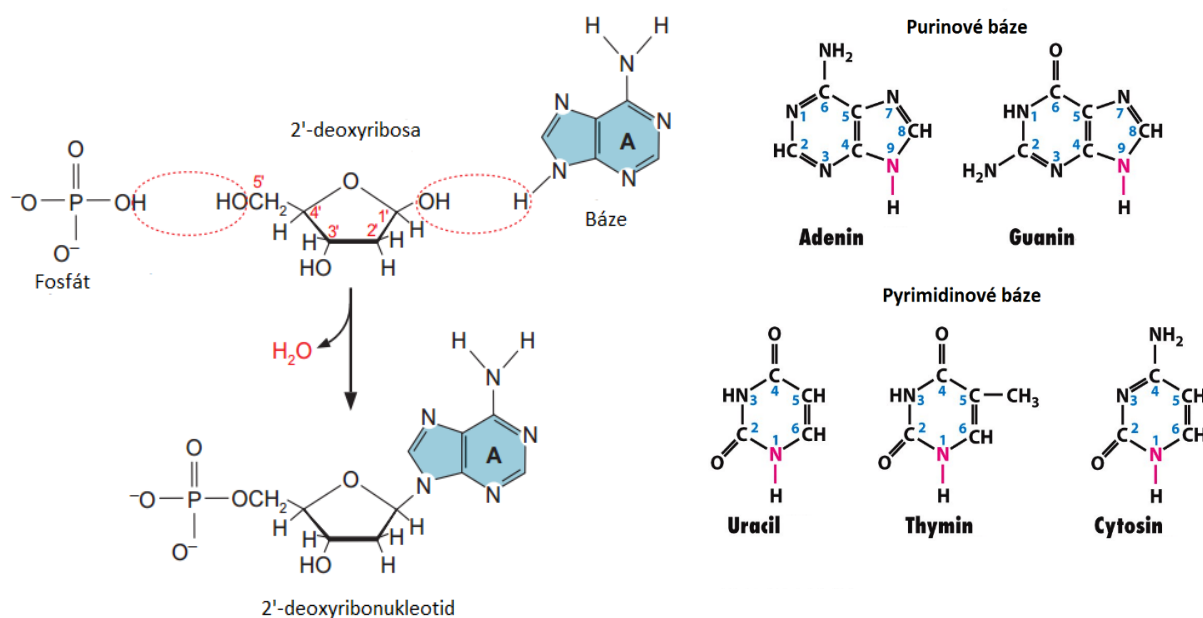
Základní stavební jednotkou nukleové kyseliny je nukleotid skládající se ze sacharidové podjednotky, ribózy nebo deoxyribózy, fosfátu, a na ribóze glykosidickou vazbou navázané báze (Obrázek 2). (Deoxy)ribóza se na fosfátovou skupinu váže tak zvanou fosfodiesterovou vazbou a tvoří tak nevětvenou cukr-fosfátovou páteř. Na takto vzniklé páteři lze definovat směr (někdy také označovaný jako polarita) a vzhledem ke stejné orientaci fosfodiesterových vazeb, fosfát vždy spojuje dvě ribózy přes 3'-OH a 5'-OH, lze jednoznačně určit 3' a 5' konce každého lineárního vlákna. V přírodě se nukleové kyseliny vyskytují buď ve formě ribonukleové, nebo deoxyribonukleové kyseliny, tyto dvě molekuly se liší pouze přítomností ribózového nebo deoxyribózového fragmentu. Majoritní báze v nukleových kyselinách jsou purinové báze adenin

a guanin a pyrimidinové báze cytosin, převážně v DNA se vyskytující thymin a v RNA se vyskytující uracil (Obrázek 2) [1,5].

### 2.2.1.1. Báze

Báze jsou planární aromatické heterocyklické sloučeniny (Obrázek 2). Všechny báze existují ve více tautomerních formách. Watson s Crickem předpokládali, že báze se vyskytují v DNA s největší pravděpodobností v jejich nejstabilnější tautomerní formě (Obrázek 2) [1,5–7]. Aromatický charakter bází způsobuje absorpci UV záření a produkci charakteristických spekter [8]. Báze mají převážně hydrofobní charakter, jsou ale schopné tvořit silné směrové vodíkové vazby [1].

Báze jsou citlivé na širokou škálu chemických látek například aminy a amidy, aromatické (polycyklické) uhlovodíky, deaminační činidla ( $\text{HNO}_2$ ) nebo alkylační reagens jako je dimethylsulfát [5]. Působení výše vyjmenovaných látek se může projevit změnou ve struktuře bází (alkylace, acylace, deaminace atd.), jiné látky se mohou interkalovat mezi báze. Tyto změny mohou vést ke tvorbě mutací [8]. V organismech se také vyskytují přirozeně modifikované báze, například 5-methylcytosin je přítomný u 5 % cytosinových nukleosidů eukaryotní DNA [1].



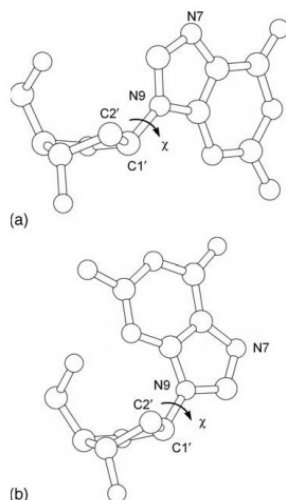
Obrázek 2: Stavební jednotkou nukleových kyselin je nukleotid. Skládá se z fosfátu, ribózy a báze. V případě DNA nalezneme v nukleotidu 2'-deoxyribózu, v RNA pak ribózu. Na obrázku vlevo je naznačen vznik dAMP kondenzací [6,9]. Vpravo jsou zobrazeny báze v nejpravděpodobnější tautomerní konformaci, převzato a upraveno z [9].

### 2.2.1.2. Ribóza

Pětičlenný kruh ribózy není planární, konformace kruhu je popsána pěti (čtyřmi nezávislými) torzními úhly [5,10]. Toto tak zvané krabatění ribózového kruhu pramení ze stereochemie tetraedrického  $sp^3$  uhlíku a nevazebných interakcí substituentů na ribózovém kruhu [11]. Energeticky nejstabilnější konformace je taková, při které jsou substituenty nejdále od sebe [5]. Bylo pozorováno mnoho konformací, nejčastěji se vychylují dva atomy z roviny definované zbylými třemi. Důležitý je směr vychýlení atomu. Pokud je směr větší výchylky na stejné straně jako je báze a vazba uhlíků  $C4'-C5'$ , pak se vychýlený atom označuje jako *endo*, pokud je na opačné straně, jedná se o konformaci *exo*. Nejčastěji se vyskytující a nejstabilnější konformace jsou  $C2'-endo$  a  $C3'-endo$  (Obrázek 3) [1,5]. Větší výchylka ve směru *endo* je kompenzována menší výchylkou ve směru *exo*. Vzhledem k různým energetickým bariérám mezi konformacemi může docházet ke změnám konformace. Na výskyt konformace ribózy má vliv i báze. Izolované nukleosidy s purinovou bází vykazují preferenci k  $C2'-endo$  a ty s pyrimidinovou  $C3'-endo$  [5].

### 2.2.2. Glykosidická vazba

Vazba spojující deoxyribózu s bází se označuje jako glykosidická, pro puriny je to vazba  $C1'-N9$  a pro pyrimidiny  $C1'-N1$ . Glykosidický torzní úhel ( $\chi$ ) je definován pro puriny jako  $O4'-C1'-N9-C4$  a pro pyrimidiny  $O4'-C1'-N1-C2$  (Obrázek 3) [5]. Vzhledem ke strukturním omezením nabývá torzní úhel  $\chi$  pouze některých hodnot. Teorie předpověděla energeticky dvě stabilní oblasti, tak zvané *syn* a *anti*. Konformace *anti* má  $N1$ ,  $C2$  purinů a  $C2$  a  $N3$  pyrimidinů namířenou směrem ven od deoxyribózy tak, že vodík na  $C8$  purinů a  $C6$  pyrimidinů je nad

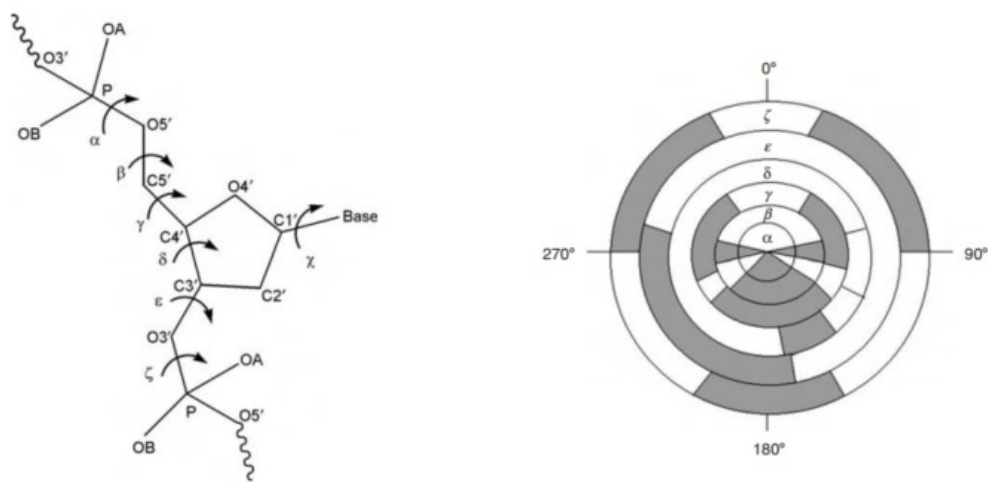


Obrázek 3: (a) Guanosinový nukleosid v *anti* konformaci. (b) Guanosinový nukleosid v *syn* konformaci, převzato a upraveno z [5].

deoxyribózovým kruhem. Konformace *syn* má tuto orientaci opačnou. Krystalové struktury potvrdily, že *syn* konformace je pro puriny méně energeticky výhodná než konformace *anti*, z důvodu menšího počtu sterických kolizí [5].

### 2.2.3. Cukr-fosfátová páteř

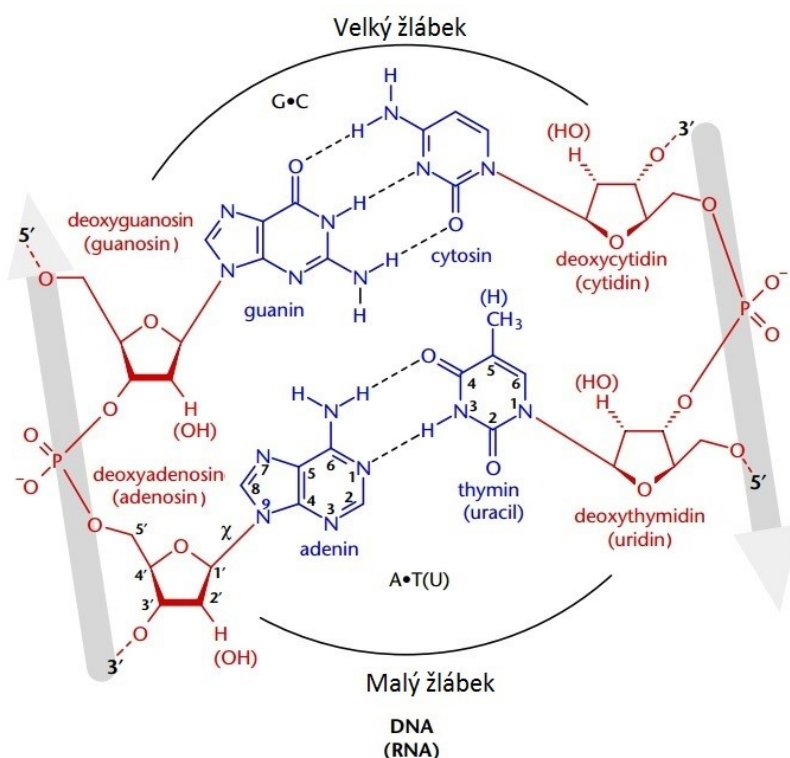
Cukr-fosfátová páteř oligonukleotidu má šest variabilních torzních úhlů ( $\alpha$ ,  $\beta$ ,  $\gamma$ ,  $\delta$ ,  $\epsilon$  a  $\zeta$ ). Hodnoty těchto úhlů spolu úzce korelují a jsou omezeny na několik diskrétních intervalů (Obrázek 4) [12].



Obrázek 4: Vlevo je zobrazeno šest torzních úhlů fosfodiesterové páteře, vpravo je konformační kruh, ten ukazuje dovolené hodnoty šesti torzních úhlů (šedá barva), převzato a upraveno z [5].

### 2.3. Párování a stabilita bází

Z experimentálních dat bylo patrné, že se molární poměry adeninu ku thyminu a cytosinu ku guaninu rovnají [13]. To vedlo k hypotéze, že páry bází A·T a C·G, označované nyní jako kanonické nebo Watsonovy-Crickovy (W-C) páry, jsou svázány specifickými vodíkovými můstky (Obrázek 5) [5,6]. Tyto páry bází mají jedinečné postavení mezi ostatními možnými kombinacemi párů. Omezení jejich párování je důsledkem isostericity obou párů, energetické stabilizace vodíkovými můstky a minimálního sterického pnutí cukr-fosfátové páteře [14]. Ačkoliv jsou báze planární, jejich párování je prostorově flexibilní. Flexibilita je ovlivněna samotnými bázemi a z velké části i bázemi nad a pod zkoumaným párem bází [5]. Nekanonické páry (jsou diskutovány dále) nebo páry v nichž je jedna z bází v *syn* orientaci glykosidického úhlu, mají silně destabilizující efekt [14].



Obrázek 5: Znázorněno je párování bází a schematicky je zakreslena pozice žlábků, převzato a upraveno z [1].

## 2.4. Sekvence a tvorba duplexu DNA

Pořadí bází na cukr-fosfátové páteři, jejich sekvence, určuje primární strukturu molekuly DNA [1,6,9]. Podle počtu nukleotidů se mohou vlákna dělit na oligonukleotidy a polynukleotidy [1,9]. Jako oligonukleotidy se označují krátká vlákna s počtem nukleotidů do 50, vlákna delší než 50 nukleotidů jsou označována jako polynukleotidy [1]. Interakce dvou komplementárních vláken vede k tvorbě duplexu DNA. Dalším důležitým faktorem ovlivňujícím stabilitu DNA jsou patrové interakce nad sebou se vyskytujícími bázemi. Patrové interakce vznikají v důsledku interakcí  $\pi$  elektronů aromatických jader [8]. Tvorba duplexu a jeho stabilita je tedy důsledkem součinnosti minimálního sterického pnutí na cukr-fosfátové páteři, isostericity párování bází a patrových interakcí mezi aromatickými jádry [14]. Silný difrakční signál s periodou 3,4 Å naměřený na vláknech DNA ukázal, že molekula DNA má tvar šroubovice s podélnou periodou v hodnotě 34 Å [2,15,16]. Z toho vyplývá, že jedna otočka šroubovice je dlouhá 34 Å, báze jsou od sebe vertikálně vzdálené právě 3,4 Å [2,15]. Tak zvaný helikální twist této formy DNA, neboli průměrná otočka na nukleotidový pár, je pak 36°.

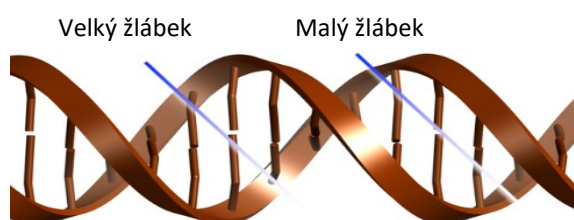
### 2.4.1. Duplex jako základní strukturní a funkční forma DNA

DNA v drtivé většině tvoří dvojitou pravotočivou šroubovici s antiparalelními vlákny [1,2,5,14]. Antiparalelní uspořádání vláken bylo potvrzeno rentgenovou krystalografií. Na každém vlákně je v jedné otáčce v průměru 10 bází, přičemž každá báze je podélně párována vodíkovými můstky s bází na druhém vlákně. Díky tomuto uspořádání se aromatická jádra bází účastní patrových interakcí a hrany bází s hydrofilními skupinami se mohou podílet na tvorbě vodíkových můstků s molekulami rozpouštědla, nízkomolekulárními látkami nebo mohou fungovat jako rozpoznávací místa pro proteiny [5,14].

Ne všechny možnosti párování jsou v duplexu DNA dovoleny ze sterických důvodů, např.: párování dvou purinů tvoří vypouklá místa v cukr-fosfátové páteři, která jsou při replikaci s velkou pravděpodobností opravena [5,6]. W-C párování má důležitý biologický význam, protože pokud se na jednom vlákně vyskytuje například adenin, musí se na druhém vlákně párovat s thyminem. Toto omezení párování je označováno jako komplementarita bází či nukleotidů [1,6]. V důsledku toho obě vlákna nesou stejnou genetickou informaci a mohou fungovat jako templát pro replikaci nebo transkripci [7].

Duplex DNA je oproti duplexu RNA relativně flexibilní, ohebný. Byly pozorovány deformace konformace v krystalových strukturách komplexů proteinů s DNA. I v silně ohnutých duplexech ale zůstává zachováno kanonické párování bází [14].

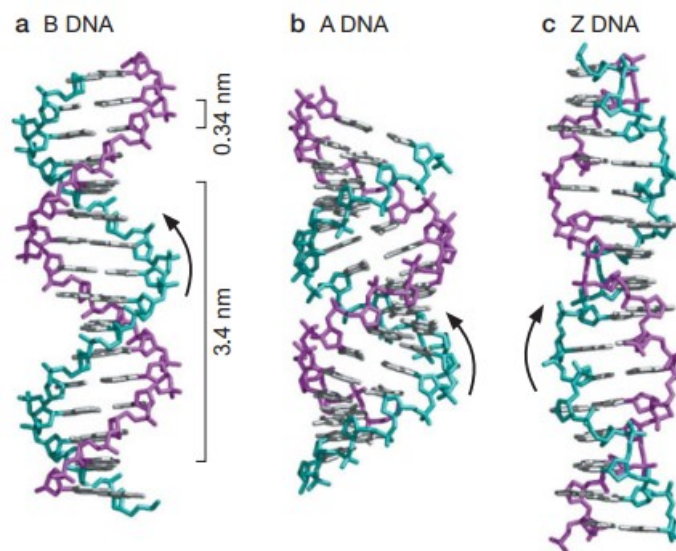
Párování bází a antiparalelní směr vláken ve dvoušroubovici má za následek, že protilehlé glykosidické vazby mezi ribóзами a bázemi nejsou kolineární, ale svírají spolu úhel menší než  $180^\circ$ . Důsledkem je tvorba tzv. žlábků (Obrázek 6). Podle jejich geometrií rozlišujeme tak zvaný velký a malý žlábek. Velký žlábek je na straně, kde jsou atomy pyrimidinových bází N4/O4, C5 a C6 a purinových N6/O6, N7 a C8 [14]. V malém žlábků jsou atomy O2 pyrimidinů a N3 purinů. Tyto atomy zároveň tvoří pomyslné dno žlábků, které, jak bylo popsáno výše, hraje důležitou roli při interakcích s ligandy [5,14].



Obrázek 6: Schematicky znázorněný velký a malý žlábek.

## 2.5. Strukturní formy duplexu

Podle rozměrů jednotlivých parametrů lze roztrždit úseky DNA do několika základních konformerů (Obrázek 6). Forma dnes známa jako B-DNA byla poprvé popsána Watsonem a Crickem [2,5]. Pokud sekvenci podrobíme podmínkám s nižší relativní vlhkostí (65-75%) získáme A-DNA, to dále dokazuje strukturní flexibilitu nukleových kyselin [1,5]. B-DNA i A-DNA jsou pravotočivé dvoušroubovice, určité sekvence jsou schopné zaujmout strukturu třetí formy DNA, Z-DNA, která je levotočivá [5,9].



Obrázek 7: Tři základní konformery DNA, zleva: B-DNA, A-DNA a Z-DNA. Šipkou je naznačen pravotočivý a levotočivý charakter konformerů, převzato a upraveno z [6].

B-DNA je dominantní formou DNA za normálních fyziologických podmínek. Průměr helixu je 20 Å. Jedné otáčky odpovídá 10 párů bází, ty jsou kolmé na osu helixu. Velký a malý žlábek mají podobnou hloubku, liší se však šířkou [8]. Velký žlábek dovoluje jednoznačné rozlišení sekvence díky přítomnosti oxo- nebo aminoskupiny na uhlících purinů C6 a pyrimidinů C4. B forma DNA se dále dělí na BI a BII formy, které se od sebe liší hodnotami torzních úhlů v cukr-fosfátové páteři [14].

B-DNA může být přeměněna na A-DNA při nízké relativní vlhkosti, například přidávkem alkoholů, nebo zvýšením koncentrace kovových iontů. Tato přeměna je reverzibilní. A-DNA je typická pro DNA-RNA hybridy a duplex RNA. Dvoušroubovice v A formě je kratší a širší než u B-DNA [8]. Ze strukturních dat plyne, že velký žlábek je úzký a hluboký a malý žlábek je široký a mělký [5]. Geometrie obou pravotočivých forem, B a A, jsou plně kompatibilní s tvorbou W-C párů.



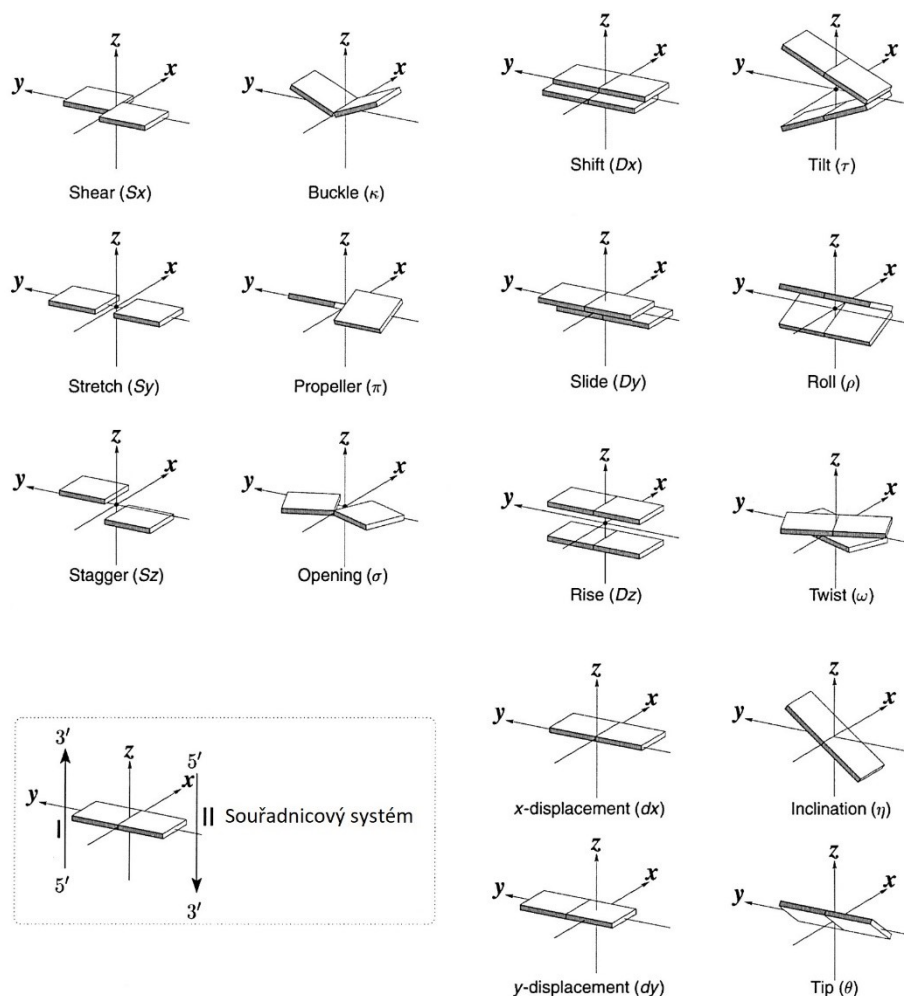
Z-DNA forma DNA je levotočivá. Vzhledem k B-DNA a A-DNA je Z-DNA relativně úzká a vysoká. Z-DNA je favorizovaná střídáním pyrimidin-purinových sekvencí, hlavně (dC-dG)<sub>n</sub>. W-C párování je v Z formě zachováno díky střídání orientací glykosidických vazeb: u cytosinu konformace *anti* a u guaninu konformace *syn*. Malý žlábek je v Z-DNA hluboký a úzký, velký žlábek téměř neznatelný [8].

Konformaci DNA lze popsat několika helikálními parametry (Tabulka 1) a parametry popisující morfologii párování samotných bází. Mezi helikální parametry patří počet párů bází na jedno otočení šroubovice, vzestup osy na jeden pár bází, definovaný jako vzdálenost mezi nad sebou se vyskytujícími bázemi, dále helikální otočení (helical twist), což je úhel pootočení bází, které jsou nad sebou. Dále se pro popis šroubovice používají rozměry malého a velkého žlábků (Tabulka 1 a Obrázek 6). Základní parametry pro popis morfologie samotných bází a jejich párů jsou na Obrázku 8 [5].

Tabulka 1: Hodnoty vybraných helikálních parametrů nejčastějších dvoušroubovicových konformací, převzato a upraveno z [5].

	Počet bp na jedno otočení šroubovice	Vzestup osy na jeden bp (Å)	Helikální twist (°)	Velký žlábek – šířka (Å)	Velký žlábek – hloubka (Å)	Malý žlábek – šířka (Å)	Malý žlábek – hloubka (Å)
A-DNA	11	2,54	32,7	2,2	13	11,1	2,6
B-DNA	10	3,38	36	11,6	8,5	6	8,2
Z-DNA	12	7,25	-30	8,8	3,7	2	13,8

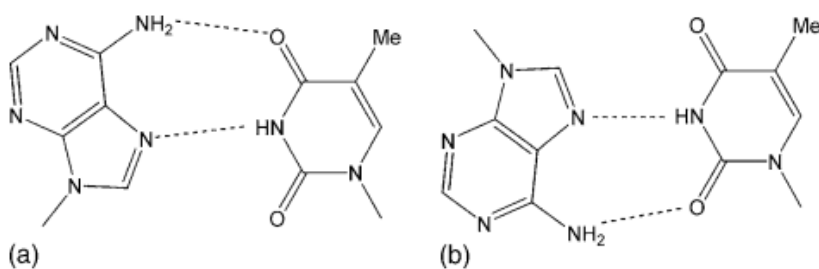
DNA je dynamická struktura. Důkazem toho je možnost přeměny mezi Z-DNA a B-DNA. Z-DNA se formuje, pokud je stabilizována vnějšími podmínkami, poté může dojít k relaxaci do B-DNA formy [8].



Obrázek 8: Parametry používané pro popis DNA, jejich názvy jsou ponechány v angličtině, převzato a upraveno z [5].

## 2.6. Další DNA struktury

Bylo pozorováno, že adenin a thymin v roztoku tvoří jiné vodíkové můstky než kanonické Watson-Crickovské párování. N7 adeninu se páruje s N3 thyminu místo N1 [1,5,8]. Toto párování se jmenuje po svém objeviteli Karstu Hoogsteenovi [5]. Hoogsteenovo párování cytosinu a guaninu je mnohem stabilnější, pokud je cytosin protonován na N3 (Obrázek 9) [8]. Díky těmto

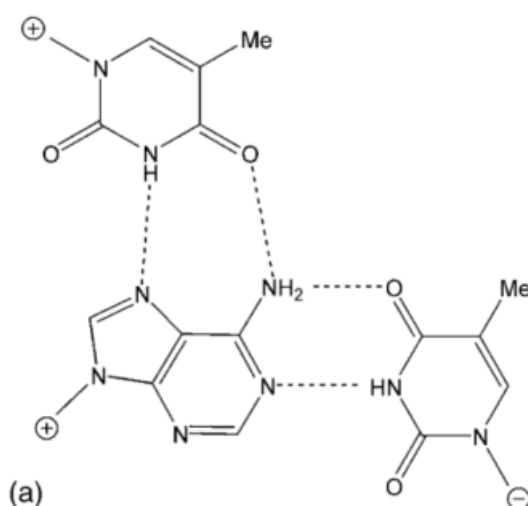


Obrázek 9: Hoogsteenovo (a) a reverzní Hoogsteenovo (b) párování, převzato a upraveno z [5].

možnostem tvorby vodíkových můstků je dosaženo velké variability při formování struktur, které je obzvláště důležité v RNA strukturách [14].

### 2.6.1. DNA trojitá šroubovice

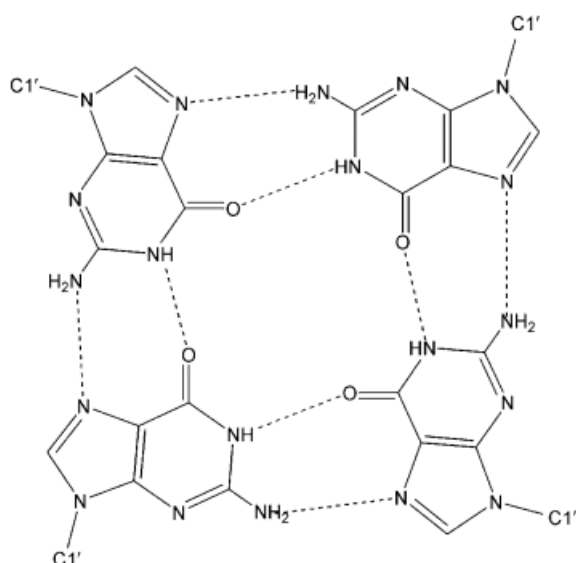
Krátce po objevení dvoušroubovice byla nejdříve postulována a následně objevena struktura složená ze dvou pyrimidinových a jednoho purinového vlákna (Obrázek 10) [1,5]. Pokud se smíchal v patřičném poměru poly(A) a poly(U), byl pozorován vznik pravotočivé trojité šroubovice, přičemž jedno párování je Watson-Crickovské a druhé je Hoogsteenovo. Podle orientace třetího vlákna se triplex může dělit na paralelní a antiparalelní. Podobnou strukturu může vytvořit i vlákno poly(G) a dva poly(C), třetí vlákno poly(C) však musí být protonováno [5,17]. Ačkoliv k protonaci cytosinu dochází při pH 5, což je výrazně pod fyziologickým pH, substituce methylu nebo bromu na C5 cytosinu vedla ke stabilizaci trojité šroubovice při pH 7 [5].



Obrázek 10: Párování T•AT v trojšroubovici, převzato a upraveno z [5].

### 2.6.2. Guaninový tetraplex

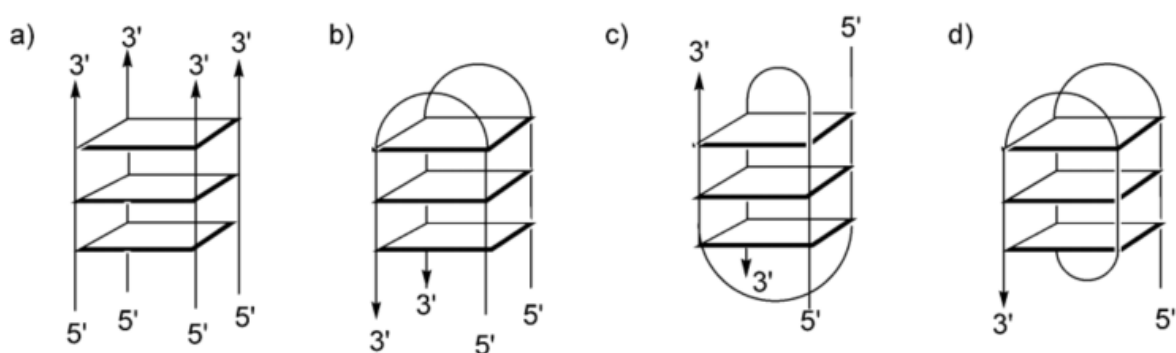
Tato velmi stabilní struktura je tvořena čtyřmi guaniny zapojenými v planární tzv. G-tetrádě (Obrázek 11). Pro strukturní stabilitu je třeba mít alespoň dvě tetrády nad sebou [5,18], struktura je dále stabilizována kovovým kationtem ležícím mezi tetrádami. Tato struktura má značný biologický význam [1,5,19]. Vyskytuje se na koncích eukaryotních chromozomů, tzv. telomerách, kde chrání DNA před nukleázami. Lidská telomerní oblast je dlouhá asi 8-10 kbp, avšak koncová oblast o velikosti 100-200 bp je jednovláknová a asociovaná s jednovláknovými vazebnými proteiny [1,5]. V případě absence těchto proteinů není tato oblast konformačně omezená a může formovat vlásenky, které jsou dále stabilizovány tvorbou tetraplexu. Tato sekvence může být také formována oligonukleotidy tvořených krátkými guanosinovými traktů rozdělenými sekvencemi jiných nukleotidů [5].



Obrázek 11: Struktura a párování čtyř guanosinů v G-tetrádě, převzato a upraveno z [5].

Několik těchto struktur vykazuje vlastnosti aptameru [5,20]. Některé uměle syntetizované oligonukleotidy vycházející z telomerní sekvence d(TG<sub>3</sub>AG) vážou silně HIV glykoproteiny [21].

Je několik možných topologií tvorby tetraplexu dané sekvence nukleotidů (Obrázek 12) [5,18]. Nukleotid tvořící guaninovou tetrádu mohou být intermolekulární nebo intramolekulární tj. tvořeny jedním nebo více vlákny [5,22]. Tyto dvě možnosti se dále liší polaritou vlákna, která mohou být paralelní nebo antiparalelní. Zdali bude vlákno paralelní nebo antiparalelní závisí na délce a sekvenci a v některých případech i na použitém kationtu, ten může být vázán v rovině tetrády na čtyři O6 kyslíky, nebo může být vázán mezi dvěma tetrádami [5].

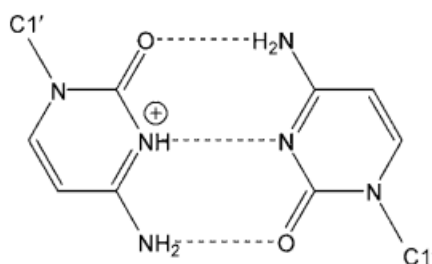


Obrázek 12: Různé struktury tetraplexu: a) paralelní tetraplex složený ze čtyř vláken, b) bimolekulární tetraplex se smyčkami na hranách, c) bimolekulární tetraplex s diagonálními smyčkami a d) příklad unimolekulárního tetraplexu, převzato a upraveno z [12].

Tetraplex má helikální twist  $36^\circ$ , tedy podobný jako dvoušroubovice DNA. Všechny tetraplexy mají čtyři žlábký ohraničené ribosafosfátovou páteří, jejich rozměry jsou závislé na celkové topologii [5].

### 2.6.3. I-Motiv

Difrakční studie poly(C) vláken ukázaly, že při mírně kyselém pH se vyskytuje  $C \cdot C^+$  párování (Obrázek 13) [1,5]. Tato struktura se vyskytuje přirozeně v telomerních sekvencích jakožto komplementární struktura ke guaninovým tetraplexům [5].



Obrázek 13: Párování cytosinu s druhým protonovaným cytosinem, převzato a upraveno z [5].

NMR studie oligonukleotidů bohatých na cytosin ukázala, že namísto duplexu se tvoří čtyřvláknový tetraplex. Základní jednotku tvoří  $C \cdot C^+$  pár, který stabilizuje dvě vlákna do duplexu, dva takové duplexy jsou následně interkalovány do podoby tetraplexu [5].

Stejně jako G-tetraplex je i-motiv slibným cílem pro terapeutický zásah.

## 2.7. Metody studia struktury nukleových kyselin

Mezi základní metody studia struktury biopolymerů patří difrakce X-paprsků na monokrystalech a spektrální metody [1,23].

### 2.7.1. Difrakce X-paprsků na monokrystalech

Jedná se o dominantní metodu určování struktury biopolymerů. Princip metody je založen na interakci X-paprsků s molekulami periodicky uspořádanými v krystalu [1,5]. Krystalizace biopolymeru představuje první kritický krok této metody. Další kroky představují krystalizaci, manipulaci s krystalem, měření difrakcí, zpracování dat, řešení struktury, její rafinaci a validaci. Vystavíme-li krystal záření o vhodné vlnové délce v řádech 1 Å dojde při splnění tak zvané Braggovy podmínky k jevu nazvanému difrakce neboli konstruktivní interference [24].

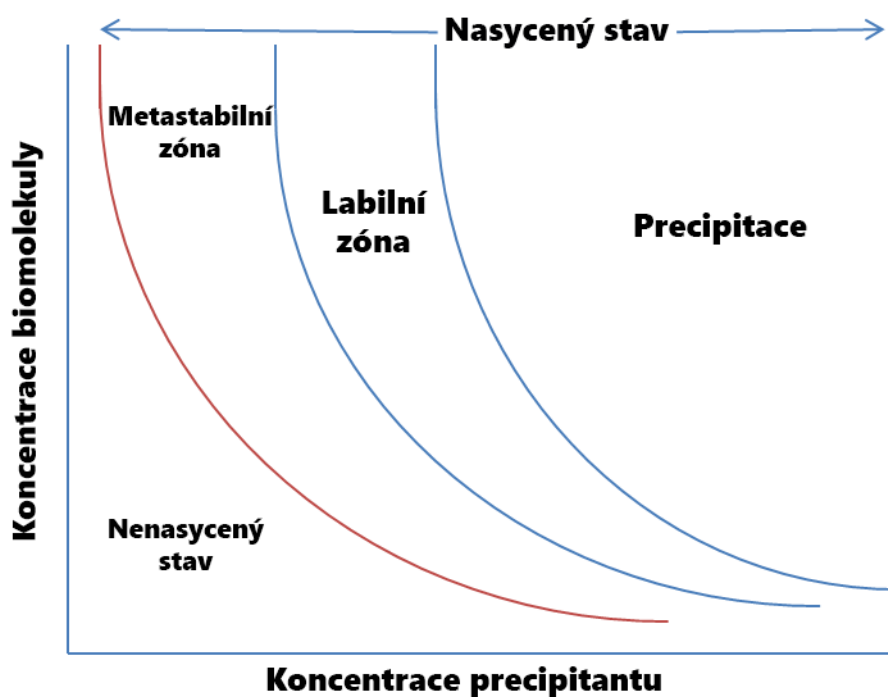
#### 2.7.1.1. Příprava materiálu

Příprava materiálu ke krystalizaci spočívá v produkci a purifikaci molekul v dostatečném množství a čistotě. Zejména při přípravě proteinů se jedná se o empirickou fázi výzkumu, kdy

se pro každý protein volí specifická kombinace expresního systému a purifikačního postupu. Pro úspěšnou krystalizaci je například potřeba koncentrace proteinu či DNA alespoň 10 mg/ml [1,23].

#### 2.7.1.2. Příprava krystalů

Existuje několik technik používaných při tvorbě krystalů, jsou však založeny na stejném principu – přechod z roztoku do pevného stavu [5,23]. Tento jev lze dokumentovat na krystalizačním fázovém diagramu na Obrázku 14 [25].

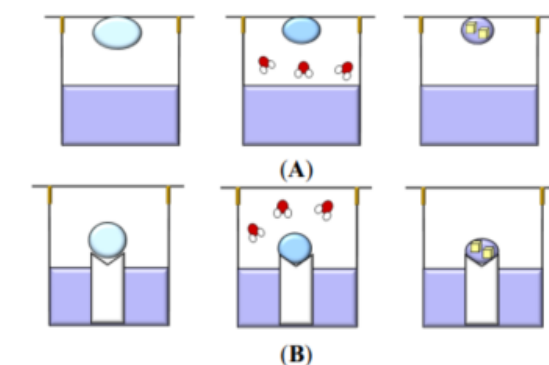


Obrázek 14: Fázový diagram krystalizace biomolekul, ukazuje rozpustnost biomolekuly (protein, DNA, RNA ...) jako funkci koncentrace precipitantu. Plochy na grafu reprezentují různé stavy roztoku, od nenasyceného stavu, po metastabilní, labilní a precipitační zónu, převzato a upraveno z [25].

Biomolekula, může se jednat o protein nebo nukleovou kyselinu, zůstává rozpuštěna v roztoku jen do určité koncentrace, po dosažení této limitní koncentrace již není roztok fázově homogenní a začíná se objevovat pevná fáze. Pokud je rychlost tvorby krystalu malá, systém přejde do metastabilního stavu (zóny ve fázovém diagramu), v tomto stavu nedochází k žádné viditelné tvorbě krystalu. V labilní zóně je rychlost tvorby krystalu dostatečná a lze pozorovat vznik krystalů. Pokud bude nasycenost roztoku ještě větší než v labilní zóně mohou se začít formovat agregáty biomolekul nebo dojde k separaci další kapalně fáze, roztok přešel do precipitační fáze [25].

Jelikož je fázový diagram funkcí mnoha proměnných – pH, koncentrace biomolekuly, druhu a koncentrace precipitantu, aditiva, teploty atd. – lze pozměněním některého z těchto parametrů výrazně ovlivnit úspěšnou přípravu krystalu [25,26].

Je několik možností experimentálního sestavení systému využívajícího difuze par. Tato technika využívá vypařování a difuze těkavých složek mezi malou kapkou (0,5 – 10  $\mu$ l) obsahující krystalizovanou biomolekulu, pufr, precipitant a aditiva a rezervoárem, který obsahuje vyšší koncentraci pufru, precipitantu a aditiv než kapka. Nejčastěji se používají techniky visící a sedící kapky, u visící kapky je kapka umístěna vespod krycího sklíčka (Obrázek 15 (A)), u sedící kapky je kapka umístěna na malý můstek nad rezervoárem (Obrázek 15 (B)) [27].



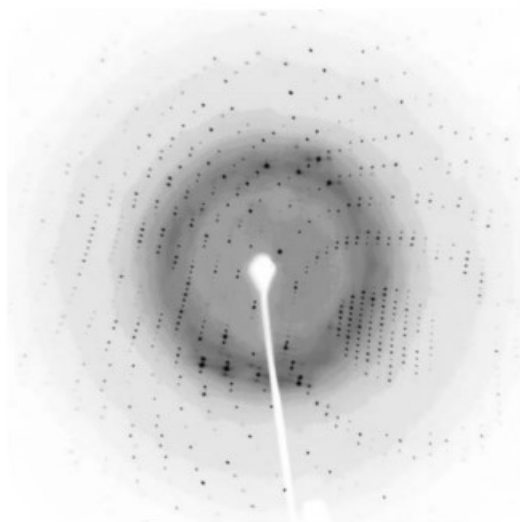
Obrázek 15: (A) sestavení experimentu při použití visící kapky a (B) sedící kapky, převzato a upraveno z [19].

Mezi další používané metody a techniky patří krystalizace dialýzou, gelová krystalizace nebo například krystalizace přes volné rozhraní [27].

### 2.7.1.3. Měření difrakce

Krystal se vyloví pomocí smyčky z kapky nad rezervoárem a podrobí se kryoprotekční ochraně ponořením do kapalného dusíku [27]. Smyčka se poté umístí na goniometr v difraktometru. Experiment se provádí při teplotě kapalného dusíku. Fokuseované a monochromatizované X-paprsky jsou namířeny na krystal a detektor detekuje reflexe [26].

Měření difrakce spočívá v naměření velkého počtu reflexí (Obrázek 16). Díky tomu, že krystaly jsou symetrické objekty, některé reflexe jsou ekvivalentní a mají proto stejnou intenzitu [26]. Průměrný počet měření jedné symetricky unikátní reflexe označujeme jako redundanci či multiplicitu. Každá reflexe je charakterizována amplitudou a fází, z měření však mohou být získány jen amplitudy, informace o fázích nikoliv [25,26].



Obrázek 16: Ukázka typického difrakčního obrazce biomolekuly, převzato a upraveno z [5].

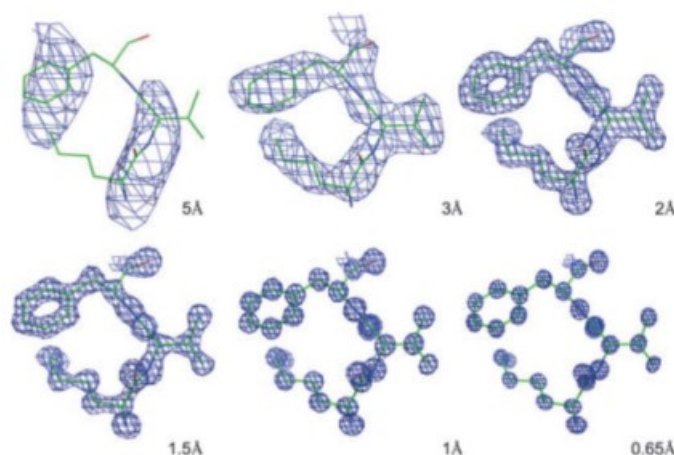
Nemožnost měření fází se nazývá fázový problém. Jeho řešení je možné různými technikami: tak zvanými přímými metodami, isomorfní záměnou, anomální disperzí, nebo molekulárním nahrazením. Přímou metodu lze použít pro malé molekuly a fáze může být získána přímo ze strukturních faktorů. Metoda se proto označuje jako přímá, protože informace o fázi byla získána z difrakčních obrazců. Isomorfní záměna spočívá v tom, že do krystalu biomolekuly se zavede látka s těžkým atomem, aniž by došlo k narušení struktury molekul a krystalové mřížce. Existují dva způsoby zavedení těžkého atomu: kovalentní vazbou (typicky brom, jód nebo selen) na specifické místo v molekule a nesespecifické namáčení (zde se používají roztoky hexaminů kobaltu, iridia, osmia či selenourey) [24]. Těžký atom difraktuje více než zbytek atomů v krystalu, poté lze přímou metodou zjistit polohu tohoto atomu v krystalové buňce a následně lze vypočítat a získat fáze reflexí. Metoda anomální disperze spočívá v tom, že schopnost rozptylovat (difraktovat) paprsky X je v blízkosti určité vlnové délky velmi citlivá na její přesnou hodnotu. Pro některé prvky (například selen a brom) je obor vlnových délek, kde k tomuto jevu tak zvané anomální difrakce dochází, vhodný pro krystalografické experimenty. Měřením na jedné nebo více vlnových délkách potom změříme proměnlivou difrakci takového. Aplikací přímých metod nebo jiné analýzy difrakčních dat (tzv. Pattersonovy funkce) zjistíme potom polohu atomu s anomální difrakcí a podobně jako u isomorfní záměny získáme hodnoty fází. Anomální signál může pocházet z atomů v nativním krystalu (atomy P, S, Cl<sup>-</sup>, K<sup>+</sup> nebo Ca<sup>2+</sup>), ty ale poskytují slabý anomální signál, nebo z vhodně voleného exogenního těžkého atomu jako zmíněný selen nebo brom. Ačkoliv nativní krystaly difrakují obvykle lépe než isomorfní krystaly, poskytují slabý signál. Poslední metoda, metoda molekulárního nahrazení je založena na



předpokladu podobnosti mezi strukturou molekuly řešené a některé již známé. Molekula se známou strukturou se použije k získání fází krystalu neznámé molekuly [24,28].

#### 2.7.1.4. Výpočet mapy elektronových hustot

Výsledkem krystalografických měření je mapa elektronové hustoty v elementární buňce (Obrázek 17). Mapa elektronových hustot může být vypočtena Fourierovou transformací z naměřených velikostí amplitud a fází reflexí ( $F_{\text{obs}}$ ,  $\Phi_{\text{calc}}$ ) [26]. Taková mapa představuje aproximaci skutečné struktury, záleží totiž na přesnosti vypočítaných fází. Jiný druh mapy, tzv. diferenční mapa, používá rozdíl naměřených a vypočítaných amplitud ( $F_{\text{obs}} - F_{\text{calc}}$ ,  $\Phi_{\text{calc}}$ ), v tomto typu mapy lze identifikovat části modelu, které byly zavedeny chybně nebo části, které chybí ale nebyly zahrnuty v modelu [26].



Obrázek 17: Příklad elektronových hustot získaných z dat s různým rozlišením, převzato a upraveno z [18].

V praxi se používá superpozice předchozích dvou map ( $2F_{\text{obs}} - F_{\text{calc}}$ ,  $\Phi_{\text{calc}}$ ), která ukazuje elektronovou hustotu a zároveň části, které vyžadují vylepšení [26].

#### 2.7.1.5. Rafinace struktury

Přesnost a spolehlivost struktury je dána částečně kvalitou měřeného krystalu, částečně kvalitou měření, které společně určují krystalografické rozlišení difrakčních dat. Správnost struktury je dána v hrubých obrysech reflexemi s nízkým a středním rozlišením a v detailech reflexemi s vysokým rozlišením [5].

K ověření správnosti struktury slouží dva základní faktory: R faktor,  $R$ , a volný R faktor,  $R_{\text{free}}$ . R faktor je definován jako suma všech pozorovaných reflexí, kde  $F_{\text{obs}}$  jsou pozorované a  $F_{\text{calc}}$  jsou vypočítané strukturní faktory, ty jsou spolu navzájem svázány vztahem (1) [5]

$$R = \frac{\sum |F_{\text{obs}}| - |F_{\text{calc}}|}{\sum |F_{\text{obs}}|}. \quad (1)$$

R faktor je často vyjádřen ve formě desetinného čísla nebo procenta. Volný R faktor se vypočítá tak, že se vezme malá část náhodně vybraných reflexí, která se nepoužije při rafinaci. Zbytek reflexí se pro rafinaci použije a  $R_{\text{free}}$  ukazuje, jak model předpovídá vynechaná data [5].

#### 2.7.1.6. Výhody a nevýhody rentgenové strukturní analýzy

Krystalografické metody jsou limitované časovou a materiálovou náročností. Celý proces od přípravy materiálu v dostatečném množství po měření difrakce, zpracování dat až po zpřesnění struktury a interpretaci výsledků trvá několik týdnů, ale i několik let [23]. Zásadní výhodou je získání definitivní struktury [5].

#### 2.7.2. Spektrální metody

Spektrální metody jsou založeny na interakci záření s hmotou v roztoku. Mezi dnes nejvíce používané patří absorpční, fluorescenční, rezonanční a chiroptické metody a pro strukturní stanovení zejména metody nukleární magnetické rezonance, NMR. Pro pomocný popis struktury se nejvíce požívají rezonanční a chiroptické metody [23].

##### 2.7.2.1. Nukleární magnetická rezonance

Principem nukleární magnetické rezonance (NMR) je detekce signálu jader s jaderným spinem v externím magnetickém poli [5,29]. Při studiu nukleových kyselin se používají oligonukleotidy obohacené o izotopy  $^{13}\text{C}$ ,  $^{15}\text{N}$  a  $^{31}\text{P}$  [1,5,8]. Dále jsou detekovatelná jádra obvyčejného vodíku, protony [5]. V základním stavu mají tyto atomy určitou orientaci svého jaderného spinu, pokud jim však dodáme energii, mohou jejich jaderné spiny zaujmout jinou orientaci, jsou v tzv. excitovaném stavu. Excitovaný stav je pro atom energeticky nevýhodný a má tendenci vrátit se do základního stavu, přitom vyzářit přebytečnou energii. Vyzářená energie se vyjadřuje jako tak zvaný chemický posun, který je definován pro každý nuklid pomocí zvolené standardní látky, jejíž rezonance se považuje za základní a posuny ostatních čar se vyjadřují jako její zlomky [23].

### 2.7.2.2. Výhody a nevýhody nukleární magnetické rezonance

NMR má výhodu oproti rentgenové strukturní analýze v tom, že je možné studovat molekuly v roztoku a je možné pozorovat změny ve struktuře spojené s kinetickými ději. Nevýhodou je stejně jako u rentgenové strukturní analýzy časová náročnost řešení a potřeba čistoty a velkého množství materiálu. NMR je dále limitována velikostí analyzované molekuly, s rostoucí velikostí roste složitost signálu [5]. NMR navíc nedosahuje přesnosti a rozlišení jako rentgenostrukturní analýza [23].

### 2.7.2.3. Chiroptické metody

Mezi nejpoužívanější chiroptické metody patří metoda cirkulárního dichroismu [23]. Vychází z interakce chirálních molekul s cirkulárně polarizovaným zářením, pro analýzu nukleových kyselin se nejčastěji používá ultrafialové záření v rozmezí 180-300 nm, při těchto vlnových délkách totiž absorbují báze nukleových kyselin. Pokud se liší absorpce pravotočivého a levotočivého záření, pak tento rozdíl se nazývá cirkulární dichroismus, lze jej kvantifikovat elipticitou,  $\Theta$ , nebo rozdílem molárních absorpčních koeficientů pravotočivého a levotočivého záření,  $\Delta\epsilon$  (2) [30]

$$\Delta\epsilon = \epsilon_P - \epsilon_L. \quad (2)$$

Ačkoliv chromofory nukleových kyselin jsou báze, které nejsou chirální, jejich cirkulární dichroismus je indukován interakcemi s okolím, jejich CD spektra jsou proto závislá na sekvenci a celkové struktuře měřené molekuly. CD spektroskopie je primárně vhodná pro měření konformačních přechodů, protože většina konformací nukleových kyselin má charakteristické CD spektrum [23,30].

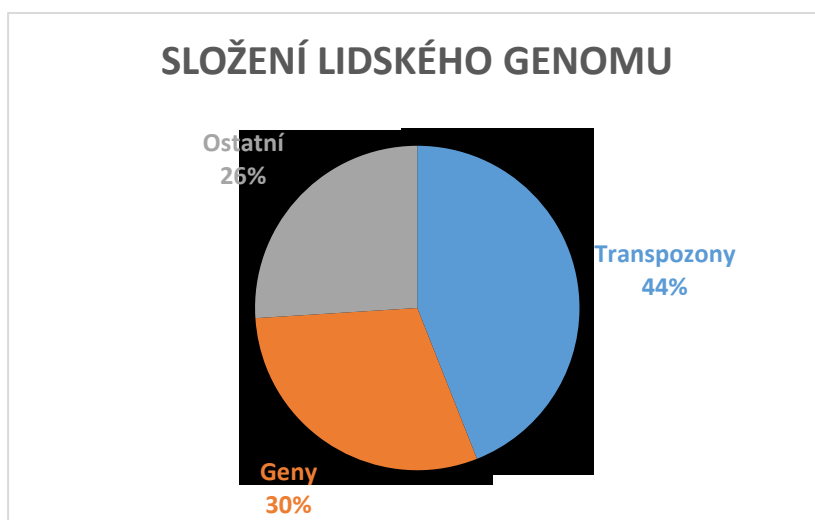
## 2.8. Strukturní databáze

Hlavním cílem strukturních databází je shromažďovat experimentálně získaná data struktur molekul [31]. Největším zdrojem strukturních dat biomolekul je strukturní databáze PDB (Protein Data Bank). Jedná se o celosvětově udržovanou databázi obsahující struktury získané pomocí difrakce paprsků X nebo neutronů, nukleární magnetické rezonance a kryo-elektronové mikroskopie [31]. V dubnu roku 2018 bylo v PDB uloženo bezmála 140 tisíc struktur, z toho téměř 90 % byly struktury získané pomocí difrakce paprsků X.

Informace o jednotlivých strukturách jsou uloženy v podobě souborů v tak zvaném PDB nebo mmCIF formátu. mmCIF je tak zvaný archivní formát. Oba formáty jsou přesně definované a obsahuje kromě informací o souřadnicích jednotlivých molekul i další položky, jako krystalografické strukturní faktory, bibliografické údaje, informace o ligandech atd.

## 2.9. Mobilní genetické elementy

Mobilní genetické elementy jsou charakterizovány svou schopností pohybu v genomu. Do této široké skupiny patří transpozony, plasmidy, bakteriofágy a introny [32]. Liší se funkcí, mechanismem pohybu v genomu i výskytem v organismu [33,34]. Mobilní genetické elementy si zasluhují velkou pozornost pro svůj potenciál v proteinovém inženýrství a pro své bohaté zastoupení v genomu, např.: asi 44 % lidského genomu je tvořeno právě transpozony (Obrázek 18) [1,35,36]. Vzhledem k obsáhlosti tématu budou dále diskutovány převážně transpozony.



Obrázek 18: V grafu je ilustrováno složení lidského genomu, upraveno a převzato z [1].

### 2.9.1. Transpozony

Transpozony jsou definované segmenty DNA schopné přenosu, díky tomu se jim někdy s nadázkou říká „skákající“ geny. Přenos DNA – transpozice – je katalyzován jedním nebo více enzymy. Přenášená DNA je transponována z lokace v genomu označované jako donor do cílové lokace nazývané akceptor. Transpozony se dělí do dvou tříd – DNA transpozony a retrotranspozony. Další možností, jak dělit transpozony, je na základě jejich schopnosti provádět transpozici samostatně. Transpozony, které obsahují gen kódující enzym schopný provést transpozici, označujeme jako autonomní. Autonomní transpozony se tedy skládají z ORF kódujícího

transponázu zpravidla ohraničeného kratšími úseky DNA sloužícími jako vazebná místa pro tuto transponázu. Neautonomní transpozony, často vzniklé působením autonomních transpozonů, potřebují pro přenos jiný (autonomní) transpozon [36,37]. Na Obrázku 19 je zobrazeno schéma bakteriálního autonomního transpozonu [37,38].



Obrázek 19: Schéma bakteriálního DNA transpozonu, IR jsou invertní repetice, převzato a upraveno [37].

#### 2.9.1.1. Retrotranspozony

Tyto transpozony (označované jako Třída 1) využívají při přenosu RNA intermediát a enzym umožňující přenos je reverzní transkriptáza [39]. Projekt HUGO ukázal, že retrotranspozony, nebo jejich pozůstatky, tvoří u člověka téměř 42 % veškeré genetické informace [35]. Díky své podstatě není třeba zavádět podtřídy, jako je tomu u DNA transpozonů, protože žádná transponáza neštěpí ani nereplikuje DNA v donorové lokaci. Tyto elementy se mohou dělit do pěti řádů LTR (zkratka z anglického „long terminal repeat“) retrotranspozony, DIRS-elementy, *Penelope* elementy (PLE), LINEs (zkratka z anglického „long interspersed nuclear elements“) a SINEs (zkratka z anglického „short interspersed nuclear elements“), toto dělení vzniklo na základě mechanismu transpozice, organizace v genomu a fylogeneze [36].

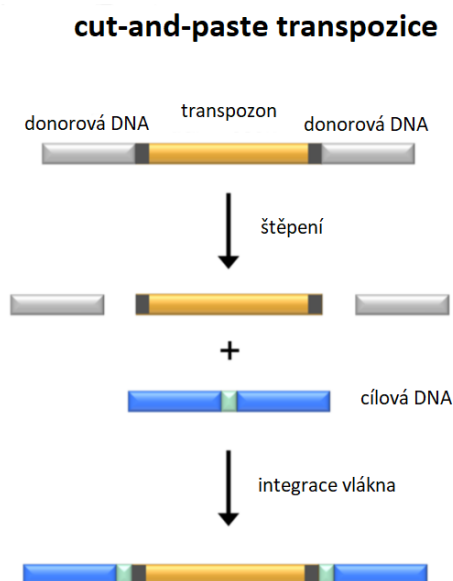
LTR retrotranspozony jsou dominantním řádem u rostlin. Jejich velikost může dosahovat až 25 kb, typicky obsahují geny *GAG* a *POL*, které kódují strukturním virům podobné proteiny, asparátovou proteázu, reverzní transkriptázu (RT), RNázu H a integrázu. DIRS-elementy obsahují místo integrázy tyrozinovou rekombinázu. PLE kódovaná RT je strukturně mnohem více podobná telomeráze než RT nalezené u LTR retrotranspozonů [36]. LINEs neobsahují LTR a byly nalezeny ve všech eukaryotních doménách. Autonomní LINEs kódují vždy alespoň RT a nukleázu. Liší se různorodostí a prevalencí v eukaryotních organismech: u člověka tvoří asi 20 % genomu v porovnání s komárem *Anopheles gambiae*, kde tvoří asi 3 % [36].

Pravděpodobnou možností vzniku SINE je náhodná retrotranspozice polymerázových transkriptů. SINE jsou to krátké elementy o délce 80-500 bp. Nejznámější SINE je *Alu* element, přítomný v lidském genomu přinejmenším v 500 000 kopiích [36].

### 2.9.1.2. DNA transpozony

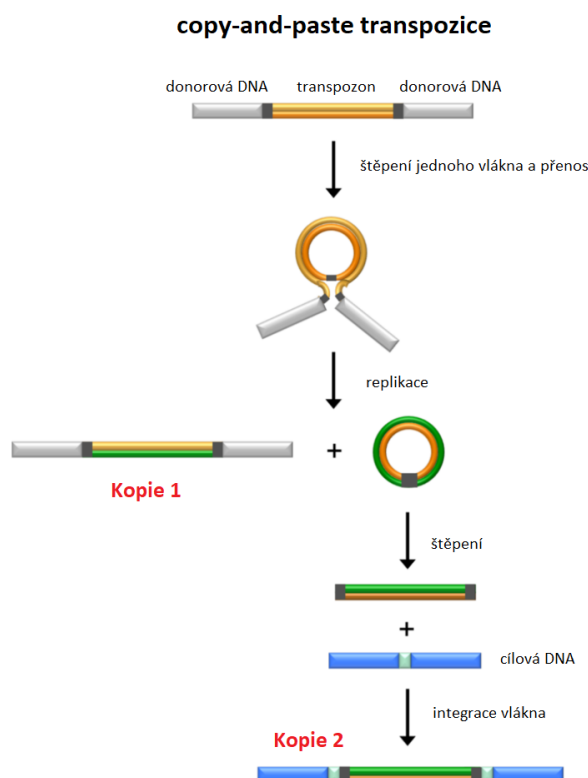
Za revoluční objev DNA transpozonů získala Barbara McClintock v roce 1983 Nobelovu cenu [40]. Tyto mobilní elementy byly nalezeny jak u prokaryotních tak eukaryotních organismů [41,42]. Jsou stále aktivní u většiny žijících organismů, u člověka, u kterého tvoří přibližně 3 % genomu, jsou inaktivované mutacemi [33,35].

Eukaryotní DNA transpozony (v literatuře označované jako Třída 2) jsou považovány za jedny z nejstarších a můžeme je dělit na podtřídy například podle toho, kolik vláken se v průběhu transpozice štěpí [36,43]. Podtřída 1 provádí transpozici mechanismem „cut-and-paste“ (Obrázek 20), je charakterizována přítomností terminálních invertních repetitiv (TIR) o různé délce. Devět podtříd je možné rozlišit pomocí TIR a duplicit vznikajících na cílovém místě (TSD – z angl. target site duplication), toho se dá využít jako diagnostického nástroje. Při transpozici této podtřídy se štěpí obě vlákna [36].



Obrázek 20: Schéma "cut-and-paste" transpozice, proces transpozice začíná štěpením donorové DNA a končí integrací vyštěpeného vlákna do cílové DNA, převzato a upraveno z [38].

Podtřída 2 provádí štěpení mechanismem „copy-and-paste“ (Obrázek 21) [36]. Tento mechanismus je někdy označován jako replikativní transpozice [36]. Do této skupiny řadíme například helitrony – ty používají při transpozici mechanismus podobný tzv. rolling-circle replikaci, z toho plyne i alternativní název rolling-circle transpozon. Tento transpozon je zajímavý tím, že kóduje protein Rep/helikázu mající endonukleázovou a 5'→3' helikázovou aktivitu [44].



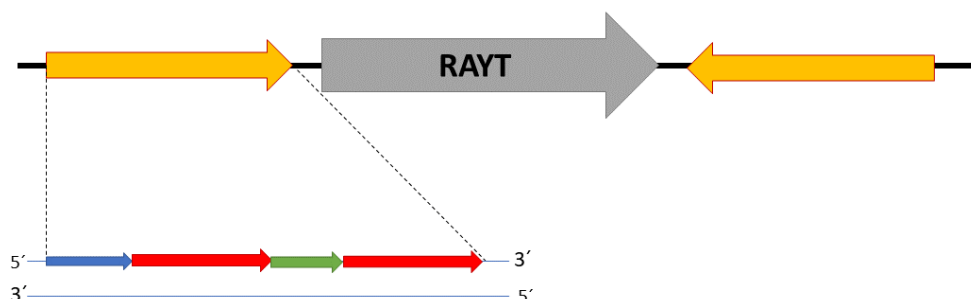
Obrázek 21: Schéma "copy-and-paste" transpozice. Začíná štěpením jednoho vlákna následované jeho přenosem, poté dochází k replikaci a vyštěpený úsek se integruje do cílové DNA. Vznikají tak dvě kopie původního transpozonu, převzato a upraveno z [38].

Prokaryotní DNA transpozony jsou označovány jako inserční sekvence (IS) [36,45]. Mohou být součástí i složitějších struktur. Jejich definice není vzhledem k rozmanitosti příliš striktní, přesto mají několik charakteristických znaků. Běžná IS je dlouhá asi 0,7 až 2,5 kb a může obsahovat i více otevřených čtecích rámců. Kromě samotné transponázy mohou IS kódovat i regulační proteiny, některé IS mohou nést tzv. pasažérský gen [46]. Rozmanitost IS kontrastuje s mechanismy jejich transpozice, které jsou ve velké míře podobné mechanismům eukaryotní transpozice [45].

### 2.9.1.3. REP – Repetitivní extragenní palindromické sekvence

REP – Repetitivní extragenní palindromické sekvence či elementy jsou nekódující IS, vyskytující se u bakterií, charakteristické přítomností evolučně konzervovaného tetranukleotidu GTAG nebo GTGG na 5' - konci, následovanou GC bohatým nedokonalým palindromem, jehož délka a sekvence jsou variabilní. Evolučně zachovaný je jejich potenciál tvořit palindromy (Obrázek 22) [47–49].

REP elementy bývají obvykle dlouhé 20-40 bp [50,51]. V bakteriálním genomu se vyskytují v několika desítkách až stovkách kopií, REP elementy se většinou vyskytují v podobě klastrů, které slouží jako vazebná místa pro proteiny [47].



Obrázek 22: Schématické zobrazení RAYT genu (šedá) ohraničeného dvěma REP elementy (žlutá). Modrou barvu je namalována rozpoznávací sekvence GT(A/G)G, červeně palindromická sekvence a zeleně nekomentární část palindromu, převzato a upraveno z [48].

REP elementy jsou mobilní, jejich transpozice, je katalyzována enzymy transponázami – RAYT (REP asociovaná tyrosinová transponáza) [48]. Tyto transponázy jsou často kódovány úseky ohraničenými REP elementy, jedná se tedy o autonomní transponázy [47]. RAYT patří do skupiny HUH nukleáz (nukleázy s HUH doménou), ty používají jeden nebo dva tyrosinové zbytky ke štěpení jednovláknové DNA [37].

Předpokládá se, že palindromické sekvence mohou formovat složitější struktury než jen vlásenky, například guaninové tetraplexy. Je možné, že REP elementy mohou tvořit i více struktur, které mohou být v rovnováze [47].

### 2.9.2. Biologický význam mobilních elementů

Bakteriální IS mají vliv na širokou škálu dějů probíhajících v buňce. Podílejí se na celogenomových editacích svou schopností expanze. Bakterie rostoucí na nutričně bohatém mediu redukuje počet exprimovaných enzymů, toho mohou docílit náhodnou transpozicí vedoucí k mutacím. IS mohou i přímo regulovat genovou expresi tím, že gen přesunou z jednoho replikonu do druhého, takto vzniklý konstrukt se označuje kompozitní transpozon [45]. Vliv eukaryotních transpozonů na architekturu genomu je podobný jako u bakteriálních IS [44].

DNA transpozice poskytuje výhodu bakteriím vystaveným antibiotikům, protože transpozony mohou přenášet geny odpovědné za rezistenci vůči antibiotikům [52]. Další nespornou výhodou



transpozonů je jejich přirozená schopnost ovlivňovat složení celého genomu a přispívat k jeho variabilitě [46].

Již na počátku výzkumu transpozonů byl rozpoznán jejich velký potenciál jakožto nástroje pro editaci genomu [52,53]. Transpozici lze využít hlavně dvěma způsoby – k zavádění náhodných mutagenezí a k implementaci genů na specifické místo [52,54]. Prvního se využívá při tvorbě knihoven či při výzkumu rakoviny, druhého, jako alternativa k systémům jako je CRISPR nebo TALEN, při zavádění genů do genomu, ačkoliv je zde vyšší šance vzniku mutace kvůli obecně malé selektivitě cílového místa [52].

### 3. Cíle práce

Cílem předkládané práce bylo:

1. Seznámit se s
  - 1.1. poznáním strukturních forem DNA
  - 1.2. experimentálními technikami stanovení topologie a struktury nukleových kyselin
  - 1.3. biochemií mobilních elementů DNA, přesněji transposonů.
2. Experimentálně studovat oligonukleotidy se sekvencemi speciálního typu transposonů tzv. REP elementů a to
  - 2.1. měřením spekter cirkulárního dichroismu a kalorimetricky
  - 2.2. krystalograficky difrakcí X-paprsků na monokrystalech.

## 4. Materiál a metody

### 4.1. Přístroje

Všechny použité přístroje jsou uvedeny v Tabulce 2.

Tabulka 2: Použité přístroje a jejich výrobce.

Přístroj	Výrobce
Stolní minicentrifuga VWR Galaxy Ministar C1413 (6000 RPM)	Thermo Scientific, USA
Mixér Grant Bio PV-1	Thermo Scientific, USA
Centrifuga Microfuge 20	Beckman Coulter, USA
Termoblok Biostep CHB-202	Thermo Scientific, USA
Váhy Ohaus Pioneer PA2102C	Thermo Scientific, USA
pH metr OrionStar A211	Thermo Scientific, USA
Magnetické míchadlo IKA RCTB S000	IKA, Německo
Automatické pipety Thermo Scientific F2	Thermo Scientific, USA
Krystalizační robot Crystal Gryphon LCP	ARI, USA
Krystalizační hotel Formulatrix RI1000	Formulatrix, USA
Chirascan Plus CD spektrometr	Applied Photophysics, UK
Bruker D8 Venture	Bruker, Německo
Synchrotron Bessy II	Helmholtz Zentrum Berlin, Německo
Kalorimetr Microcal iTC200	Malvern, UK

### 4.2. Chemikálie

Komerčně připravené použité krystalizační podmínky jsou uvedeny v Tabulce 3.

Tabulka 3: Krystalizační screenovací testy.

Screenovací test	Výrobce
Nucleic Acid Mini Screen	Hampton Research, USA
Natrix	Hampton Research, USA
Nucleix	Qiagen, Německo
Index	Hampton Research, USA

Všechny chemikálie použité pro optimalizaci byly zakoupeny u Sigma Aldrich, v kvalitě  $\geq 98\%$ .

Oligonukleotidy pro kalorimetrická měření a pro měření spekter cirkulárního dichroismu byly rozpuštěny v pufru 1 obsahujícím 59,8mM NaCl, 20mM Na<sub>2</sub>HPO<sub>4</sub>, 0,1mM Na<sub>2</sub>-EDTA a pufru 2 obsahujícím 79,8mM NaCl, 20mM NaH<sub>2</sub>PO<sub>4</sub> a 0,1mM Na<sub>2</sub>-EDTA. Výsledný pufr byl připraven smícháním pufru 1 a pufru 2 tak, aby jeho pH bylo 7,4.

Oligonukleotidy určené ke krystalizaci byly zakoupeny u Generi Biotech, byly standardně purifikovány (zbaveny přebytků solí a nedostatečně syntetizovaných oligonukleotidů) a jejich látkové množství bylo škálováno na 200 nmol. U oligonukleotidů Chom22-FL a Hpar1-FL byl pro fázování zaměněn thymin v pozici 9 za bromouracil.

Tabulka 4: Sekvence a označení krystalovaných a měřených oligonukleotidů.

Označení oligonukleotidu	Sekvence (ve směru 5' → 3')
Chom22	GTAGTGGGGCTTGCCCCACC
Chom22-FL	GTGGGGCTTGCCCCACC
Chom22-FL(9BrU)	GTGGGGCT(BrU)GCCCCACC
Chom22-11mer	GTAGGGTGGGG
Chom22-9mer	AGGGTGGGG
Hpar1	GTAGTGGGTCTTGACCCACC
Hpar1-FL	GGTGGGTCTTGACCCACC
Hpar1-FL(9BrU)	GGTGGGTC(BrU)TGACCCACC
Hpar1-16mer	GTAGGGTGGGTCTTGA
Hpar1-12mer	GGTGGGTCTTGA

Sekvence Chom22 a Hpar1 byly vybrány bioinformatickými metodami z bakterií *Cardiobacterium hominis* (Chom) a *Haemophilus parasuis* (Hpar).

Přehled škálování a purifikací oligonukleotidů pro jednotlivé metody je uveden v Tabulce 5.

Tabulka 5: Přehled škálování a druhů purifikací zkoumaných oligonukleotidů.

Metoda	Škálování (nmol)	Purifikace
Krystalizace	200	Standartní
Měření CD spekter	200	Standartní
Kalorimetrická měření	1000	HPLC

### 4.3. Postup při přípravě krystalů oligonukleotidů

Oligonukleotidy byly krystalizovány s pomocí krystalizačního robotu a uchovávány v krystalizačním hotelu a pro další krystalizaci vhodné podmínky byly optimalizovány ve 24 jamkových destičkách.

#### 4.3.1. Příprava oligonukleotidů na krystalizaci

Roztoky DNA oligonukleotidů se připravily rozpuštěním lyofilizátů na koncentraci 1,5 mM sterilní destilovanou vodou, byly uchovávány při teplotě -20 °C. Před použitím byly nejprve rozmrazeny při laboratorní teplotě 21 °C a následně byly v termobloku zahřáty na 95 °C po dobu 10 minut. Poté byly vyjmuty z termobloku a ponechány vychladnout při laboratorní teplotě. Před krystalizací byly mikrozkušavky s roztoky stočeny a promíchány.

#### 4.3.2. Tvorba krystalů s použitím krystalizačního robotu a hotelu

Ze zásobních roztoků oligonukleotidů se odebralo 80 µl do 0,2 ml mikrozkušavky. Ta se umístila do příslušného stojanu v krystalizačním robotu.

Přístroj pipetoval 70 µl zakoupené předpřipravené krystalizační podmínky do jamky v 96 jamkové destičce a poté do tří jamek nad rezervoárem pipetoval roztok DNA oligonukleotidu a rezervoáru v poměru 2:1; 1:1 a 1:2 do celkového objemu 0,3 µl. 96 jamková destička je v uskušení sedící kapky.

Poté byla destička zalepena fólií, aby nedocházelo k vysychání kapek, a umístěna do krystalizačního hotelu. V něm se po definovaných intervalech fotily jednotlivé jamky a byla sledována případná tvorba krystalů.

#### 4.3.3. Optimalizace krystalů a fázování

Podmínky, ve kterých došlo ke tvorbě krystalu, byly optimalizovány ve 24 jamkových krystalizačních destičkách, využívající sestavení visící kapky. Objem rezervoáru byl 0,5 ml. Kapka nad rezervoárem byla míchána ve stejném poměru jako v hotelu, do celkového objemu 3 µl. Připravené destičky se umístily do 18 °C.

Pro fázování bylo použito oligonukleotidy s bromouracilem. Oligonukleotidy s bromouracilem byly nasazeny do stejných podmínek, ve kterých krystalovaly původní krystaly.

#### 4.3.4. Měření krystalů

Optimalizované krystaly byly měřeny na difraktometru Bruker D8 Venture a krystaly s inkorporovaným bromouracilem na synchrotronu Bessy II (Beamline 14.2).

Bruker D8 Venture je vybavena galiovým zdrojem s vlnovou délkou 1,34 Å a detektorem Photon II. Data nativních krystalů se sbírala po 0,1° krocích s dobou 0,1 s. Krystaly s bromouracilem byly měřeny na synchrotronu Bessy II. Data se sbírala po 0,1° krocích, doba expozice byla 0,5 s. Beamline 14.2 je vybavena detektorem Dectris Pilatus 2M. Vlnová délka zdroje byla 0,9184 Å.

#### 4.4. Postup při měření spekter cirkulárního dichroismu

Roztoky oligonukleotidů byly před měřením zahřáty na 95 °C po dobu 5 minut a nechaly se vychladnout při laboratorní teplotě. Koncentrace roztoků byla 5 μM.

Spektra byla naměřena na spektropolarimetru Chirascan Plus. Měření byly vlnové délky v rozmezí od 205 do 340 nm po 1 nm kroku. Použita byla křemenná kyveta s délkou optické dráhy 1 mm. CD signál byl vyjádřen jako závislost elipticity na vlnové délce.

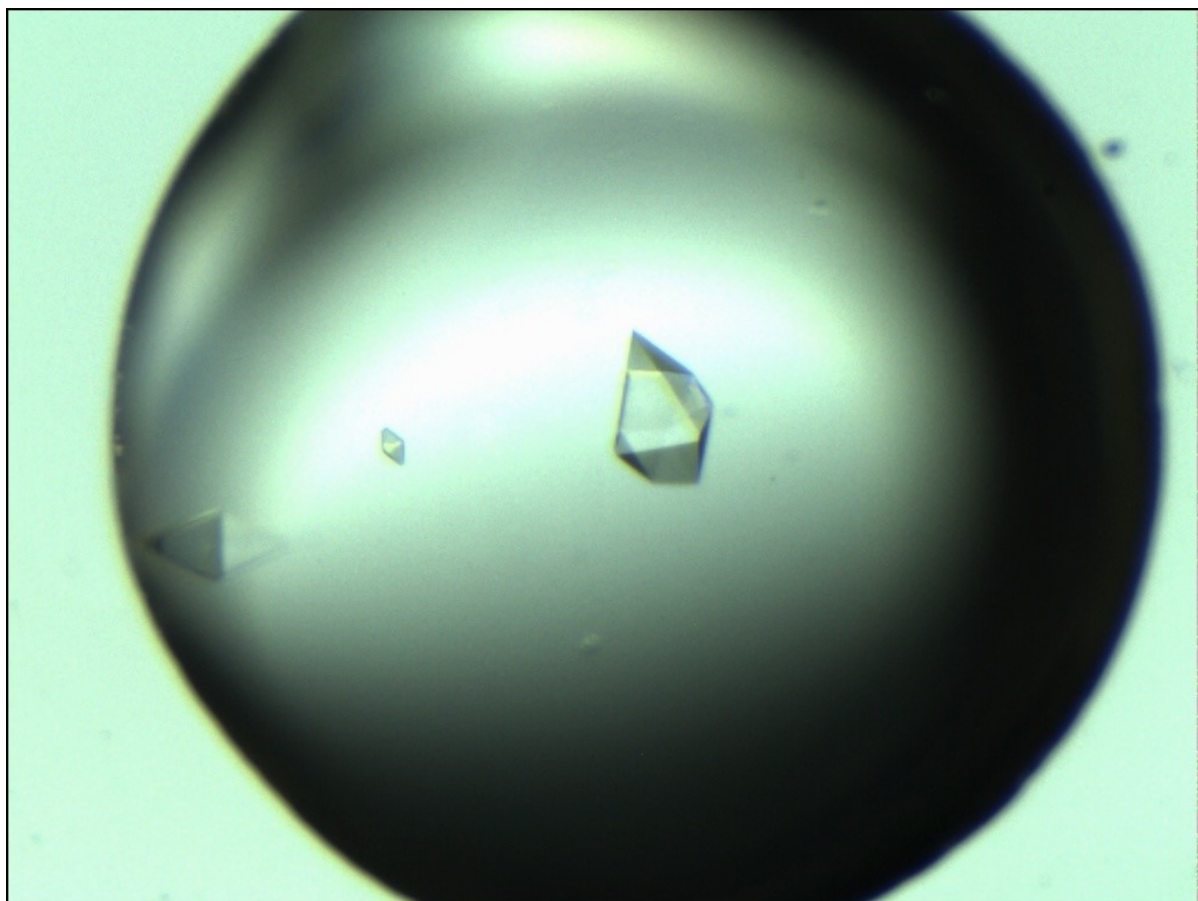
#### 4.5. Postup při kalorimetrickém měření

Měřena byla změna tepelné kapacity v závislosti na teplotě. Oligonukleotidy byly rozpuštěny na výslednou koncentraci 400 μM. Před měřením byly zahřáty v termobloku na 95 °C po dobu 5 minut. Roztok oligonukleotidu byl umístěn v měrné kalorimetrické cele, v referenční cele byl čistý pufr. Měření se provádělo v krocích po 1 °C za minutu v intervalu teplot od 5 do 105 °C na kalorimetru Microcal iTC200. Data byla statisticky zpracována v programu Origin 7.

## 5. Výsledky

### 5.1. Analýza krystalů

Podařilo se optimalizovat přípravu krystalů čtyř oligonukleotidů, Chom22-FL a jeho bromované varianty Chom22-FL(9BrU) (Obrázek 23) a Hpar1-FL a jeho bromované varianty Hpar1-FL(9BrU), jejich sekvence jsou uvedeny v Tabulce 4. Krystaly oligonukleotidů s nativními i bromovanými sekvencemi difraktovaly do krystalografického rozlišení 2,7 Å.



Obrázek 23: Vyfocený krystal oligonukleotidu Chom22-FL.

Chom22-FL krystaloval v podmínce 8 testovací sady Nucleic Acid Mini Screen obsahující 10% MPD; 0,04M trihydrát kakodylátu sodného (pH 6,0); 0,012M spermin tetrahydrochlorid a 0,08M chlorid sodný. Tyto podmínky byly následně optimalizovány dle schématu uvedeného v Tabulce 5.

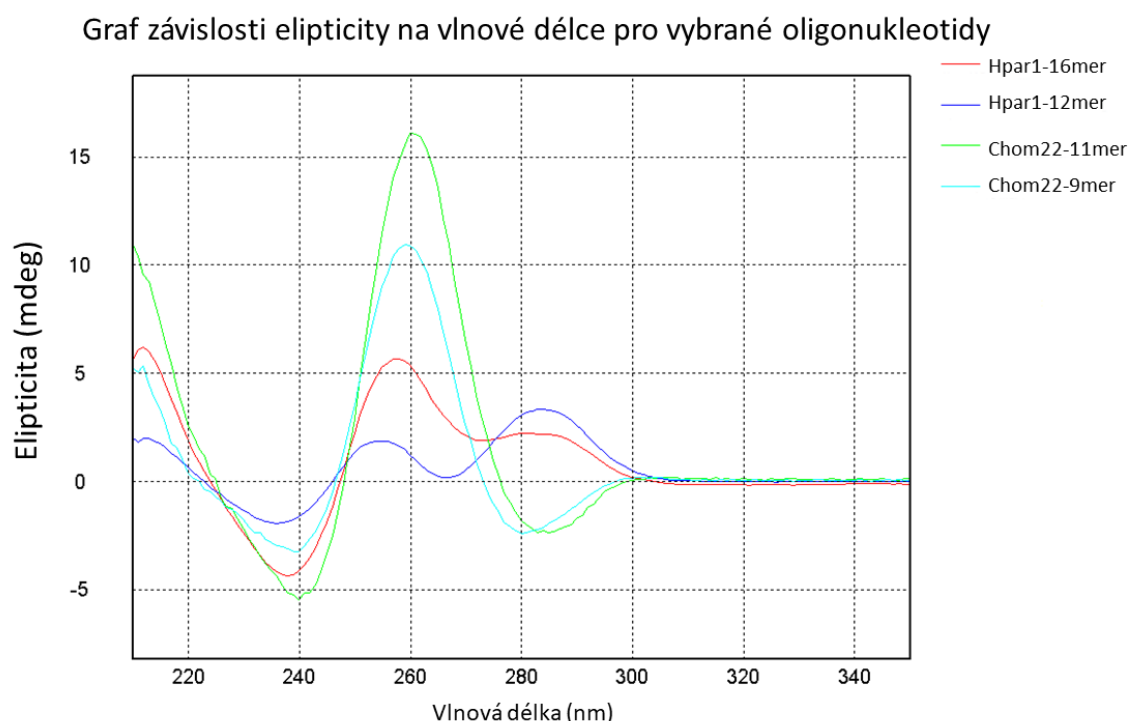
Tabulka 6: Optimalizování podmínky 8 z Nucleic Acid Mini Screen testu ve 24 jamkové destičce.

		MPD					
		5%	10%	15%	20%	25%	30%
NaCl	0,04M	0,04M Trihydrát kakodylátu sodného (pH 6,0); 0,012M Spermin					
	0,08M						
	0,12M						
	0,16M						

Obdobným způsobem se postupovalo u oligonukleotidu Hpar1-FL. Ten krystaloval v podmínce G9 z testovací sady Index. Tato podmínka obsahovala 25 % MPD; 0,04M trihydrát kakodylátu sodného (pH 7,0); 0,04M chlorid lithný; 0,08M hexahydrát chloridu strontnatého a 0,012M spermin tetrahydrochlorid. Oba oligonukleotidy, stejně jako jejich varianty s bromouracilem, vytvořily krystaly během dvou až tří dnů. Fázový problém je v současné době řešen analýzou anomální disperze atomů bromu ve dvou oligonukleotidech.

## 5.2. Spektra cirkulárního dichroismu

Na Obrázku 24 jsou zobrazeny závislosti elipticity na vlnové délce pro oligonukleotidy Hpar1-16mer, Hpar1-12mer, Chom22-11mer a Chom22-9mer.



Obrázek 24: Na obrázku je znázorněna závislost elipticity na vlnové délce pro oligonukleotidy Chom22-11mer, Chom22-9mer, Hpar1-16mer a Hpar1-12mer.

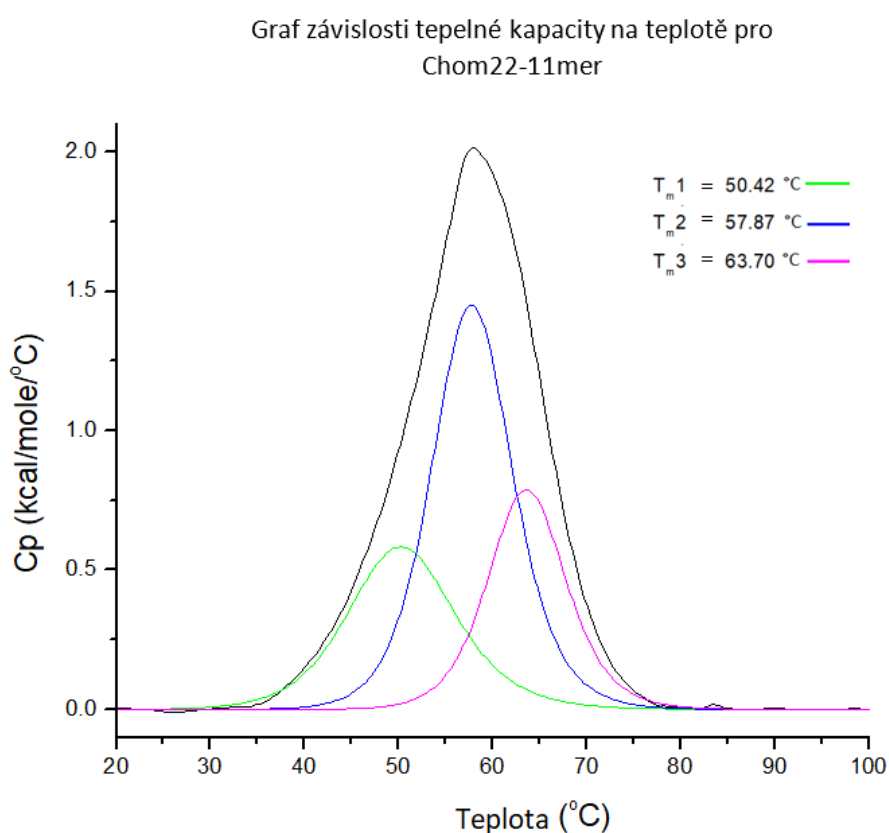
Ze spekter cirkulárního dichroismu lze vyčíst hodnoty charakteristické pro guaninový tetraplex, tj. kladná hodnota elipticity v oblasti kolem 210 a 260 nm u všech čtyř měřených oligonukleotidů. CD spektra oligonukleotidů Chom22-FL a Hpar1-FL byla změřena a publikována dříve [47] a vykazovala také charakteristické hodnoty pro guaninový tetraplex.



Výrazně vysoký vrchol při 260 nm u oligonukleotidů Chom22-11mer a Chom22-9mer naznačuje přítomnost paralelního tetraplexu. CD spektra oligonukleotidů Hpar1-12mer a Hpar1-16mer vykazují kladné hodnoty vrcholu u 290 nm. To indikuje přítomnost antiparalelního tetraplexu. Záporný vrchol v okolí 240 nm u spekter všech oligonukleotidů může značit přítomnost paralelního tetraplexu (Chom22-11mer a Chom22-9mer) nebo formování B-DNA [30].

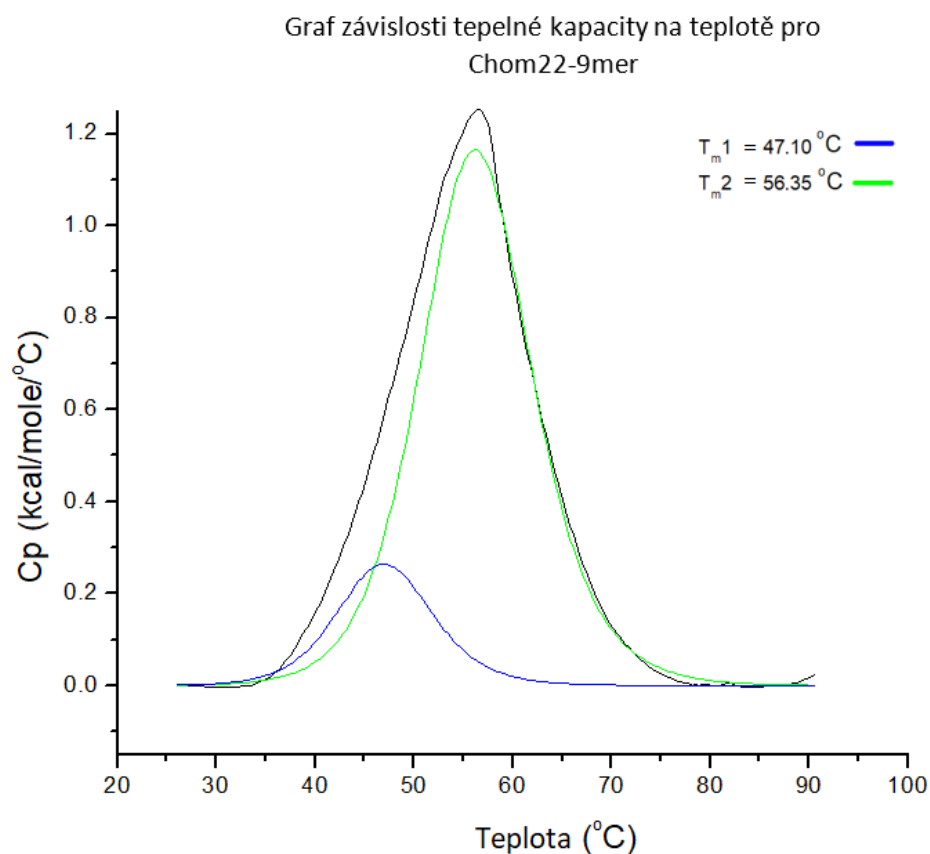
### 5.3. Kalorimetrická měření

Na Obrázcích 25-28 jsou grafy závislosti tepelné kapacity za konstantního tlaku  $c_p$  na teplotě pro oligonukleotidy Chom22-11mer, Chom22-9mer, Hpar1-16mer a Hpar1-12mer.



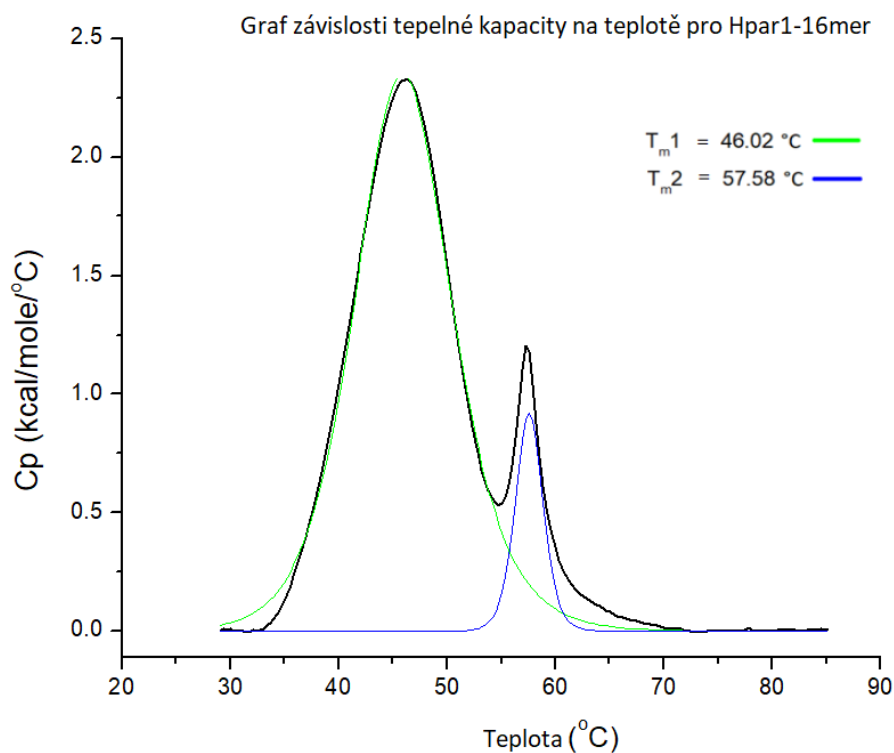
Obrázek 25: Graf závislosti tepelné kapacity na teplotě pro oligonukleotid Chom22-11mer. Černá křivka ukazuje naměřená data. Zeleně, modře a růžově jsou znázorněny dílčí křivky tání zjištěné statistickou analýzou. Vrcholům křivek odpovídají teploty tání  $T_{m1}$ ,  $T_{m2}$  a  $T_{m3}$ .

Z Obrázku 25 lze statistickou analýzou určit tři různé teploty tání ( $T_{m1} = 50,42\text{ }^{\circ}\text{C}$ ;  $T_{m2} = 57,87\text{ }^{\circ}\text{C}$  a  $T_{m3} = 63,70\text{ }^{\circ}\text{C}$ ), tedy teploty, při kterých je denaturováno právě 50 % molekul dané konformace, oligonukleotidu Chom22-11mer. Původně změřená křivka s jediným maximem (černá křivka) byla následnou statistickou analýzou metodou zjištěno, že křivka je složena ze tří dílčích křivek, kdy každá představuje křivku tání jednoho molekulárního individua. Podobná situace nastala u oligonukleotidu Chom22-9mer, avšak zde byly pozorovány jen dvě dílčí křivky, tedy dvě teploty tání:  $T_{m1} = 56,35\text{ }^{\circ}\text{C}$  a  $T_{m2} = 47,10\text{ }^{\circ}\text{C}$  (Obrázek 26).

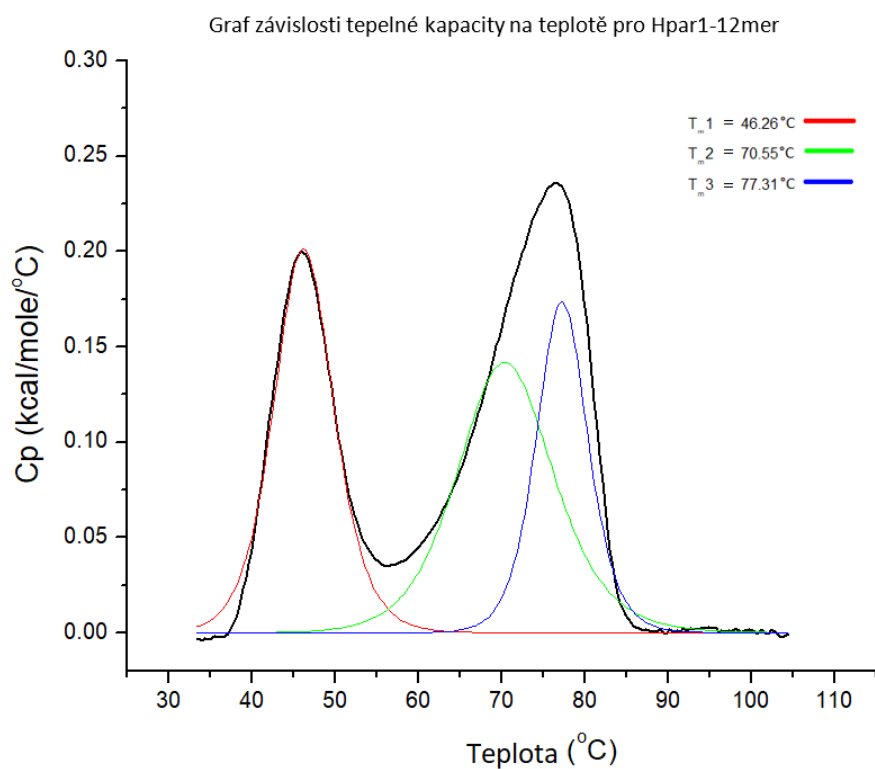


Obrázek 26: Graf závislosti tepelné kapacity na teplotě pro oligonukleotid Chom22-9mer. Zeleň a modře jsou zaznamenány dílčí křivky tání.

Na Obrázcích 27 a 28 jsou zaznamenány průběhy tepelných denurací roztoků oligonukleotidů Hpar1-16mer a Hpar1-12mer. Oproti dvěma variantám Chom22 jsou na denaturační křivce u sekvence Hpar1-16mer jasné dva vrcholy, kterým odpovídají dvě teploty tání. Pro Hpar1-16mer to jsou  $T_{m1} = 46,02\text{ }^{\circ}\text{C}$  a  $T_{m2} = 57,58\text{ }^{\circ}\text{C}$ . Denaturační křivka oligonukleotidu Hpar1-12mer je složena ze tří dílčích křivek, kterým odpovídají teploty tání  $T_{m1} = 46,26\text{ }^{\circ}\text{C}$ ,  $T_{m2} = 70,55\text{ }^{\circ}\text{C}$  a  $T_{m3} = 77,31\text{ }^{\circ}\text{C}$ .



Obrázek 27: Graf závislosti tepelné kapacity na teplotě oligonukleotidu Hpar1-16mer. Zeleně a modře jsou zaznamenány dílčí křivky tání.



Obrázek 28: Graf závislosti tepelné kapacity na teplotě pro oligonukleotid Hpar1-12mer. Červeně, zeleně a modře jsou zaznamenány dílčí křivky tání.

## 6. Diskuze a závěr

Naměřená spektra cirkulárního dichroismu (Obrázek 24) naznačují, že všechny čtyři sekvence tvoří topologicky různé guaninové tetraplexy: paralelní tetraplex pro sekvence Chom22-9mer a Chom22-11mer a antiparalelní tetraplex pro Hpar1-12mer a Hpar1-16mer. Vzhledem k poměrně vysokému počtu guaninů v sekvencích je toto pozorování očekávatelné. Je třeba mít na paměti, že interpretace spekter cirkulárního dichroismu není jednoznačná a v principu nemůže odhalit, zdali jsou v roztoku přítomny i jiné konformace s nízkou elipticitou než ty, jejichž typická spektra jsou zachycena na Obrázku 24.

Z tohoto důvodu jsme provedli kalorimetrická měření, která měla naznačit přítomnost různých konformací jednotlivých oligonukleotidů. Ze všech grafů závislosti tepelné kapacity na teplotě (Obrázky 25-28) vyplývá, že v roztoku je přítomno více konformací. Tato hypotéza je založena na faktu, že statistická analýza denaturačních křivek ukázala vždy alespoň dvě teploty tání. U oligonukleotidů Chom22-11mer a Hpar1-12mer byly pozorovány tři a u oligonukleotidů Chom22-9mer a Hpar1-16mer dvě teploty tání. Teploty tání  $T_{m2}$  a  $T_{m3}$  u sekvencí Chom22-11mer a Hpar1-12mer jsou relativně blízké, to může znamenat, že konformace, kterým tyto hodnoty odpovídají, jsou strukturně podobné.

Pro tvorbu krystalu je ale nutné mít konformačně homogenní molekulu, pokud se tedy krystaly tvoří může to znamenat, že se v koncentrovaném roztoku podporuje tvorba jedné z konformací. Je zajímavé pozorovat, že podobné sekvence Chom22-FL a Hpar1-FL tvoří krystaly ve velmi podobných podmínkách, které se liší jen přítomností kationtu ( $\text{Li}^+$  a  $\text{Mg}^{2+}$ ), který ale může hrát důležitou roli při formování výsledné struktury. Fázový problém je řešen metodou anomální disperze na atomu bromu v bromouracilových derivátech obou sekvencí.

Kratší verze sekvencí Chom22-11mer, Chom22-9mer, Hpar1-16mer a Hpar1-12mer, stejně jako původní sekvence Chom22 a Hpar1, se zatím nepodařilo vykrystalizovat. Možným důvodem je zřejmě konformační heterogenita oligonukleotidů, jak dokládají výstupy kalorimetrických měření. Nabízí se možnost vyzkoušet jinou metodu krystalizace, například použít místo krystalizace ve visící kapce krystalizaci v kapce sedící. Naměřená krystalografická data budou zpracována a struktury později publikovány.

Podařilo se splnit všechny vytyčené cíle.

## 7. Zdroje informací

1. Nelson, D.L., Cox, M.M. (2012) *Lehninger Principles of Biochemistry* (W.H. Freeman, New York) 6 edition.
2. Watson, J.D., Crick, F.F.H. (1953) Molecular Structure of Nucleic Acids. *Nature* **1953**, 737–38.
3. Wilkins, M.H.F. (1953) Molecular Structure of Deoxypentose Nucleic Acids. *Nature* **1953**, 738–40.
4. The Nobel Prize in Physiology or Medicine 1962 URL: [https://www.nobelprize.org/nobel\\_prizes/medicine/laureates/1962/](https://www.nobelprize.org/nobel_prizes/medicine/laureates/1962/) [Accessed April 16, 2017].
5. Neidle, S. (2002) *Nucleic Acid Structure and Recognition* (Oxford University Press).
6. Watson, J.D., Baker, T.A., Bell, S.P., Gann, A., Levine, M., Losick, R. (2013) *Molecular Biology of the Gene* (Benjamin-Cummings Publishing Company).
7. Watson, J.D., Crick, F.F.H. Genetical Implications of the Structure of Deoxyribonucleic Acids. *Nature* **1953**, 964–67.
8. Creighton, T.E. (2011) *The Biophysical Chemistry of Proteins* (Helvetian Press).
9. Lodish, H. (2004) *Molecular Cell Biology* (W. H. Freeman).
10. Williams, A.A., Darwanto, A., Theruvathu, J.A., Burdzy, A., Neidigh, J.W., Sowers, L.C. (2009) The Impact of Sugar Pucker on Base Pair and Mismatch Stability. *Biochemistry (Mosc.)* **48**, 11994.
11. Harvey, S., Prabhakaran, M. (1986) Ribose puckering: Structure, dynamics, energetics, and the pseudorotation cycle. *J. Am. Chem. Soc. - J AM CHEM SOC* **108**.
12. Schneider, B., Božíková, P., Nečasová, I., Čech, P., Svozil, D., Černý, J. (2018) A DNA structural alphabet provides new insight into DNA flexibility. *Acta Crystallogr. Sect. D* **74**, 52–64.
13. Chargaff, E., Lipshitz, R., Green, C. (1952) Composition of the desoxypentose nucleic acids of four genera of sea-urchin. *J. Biol. Chem.* **195**, 155–60.
14. Zvárová, J., Mazura, I. (2012) *Metody molekulární biologie a bioinformatiky* (Karolinum).
15. Franklin, R.E., Gosling, R.G. (1953) Molecular Configuration in Sodium Thymonucleate. *Nature* **1953**, 740–41.
16. Franklin, R.E., Gosling, R.G. (1953) Evidence for 2-chain Helix in Crystalline Structure of Sodium Deoxyribonucleate. *Nature* **1953**, 156–57.
17. Jain, A., Wang, G., Vasquez, K.M. (2008) DNA Triple Helices: biological consequences and therapeutic potential. *Biochimie* **90**, 1117–30.
18. Davis, J.T. (2004) 40 Jahre G-Quartetts: von 5'-GMP zur Molekularbiologie und Supramolekularen Chemie. *Angew. Chem.* **116**, 684–716.
19. Fry, M. (2007) Tetraplex DNA and its interacting proteins. *Front. Biosci. J. Virtual Libr.* **12**, 4336–51.
20. Tucker, W.O., Shum, K.T., Tanner, J.A. (2012) G-quadruplex DNA aptamers and their ligands: structure, function and application. *Curr. Pharm. Des.* **18**, 2014–26.
21. Romanucci, V., Gaglione, M., Messere, A., Potenza, N., Zarrelli, A., Noppen, S., Liekens, S., Balzarini, J., Di Fabio, G. (2015) Hairpin oligonucleotides forming G-quadruplexes: New aptamers with anti-HIV activity. *Eur. J. Med. Chem.* **89**, 51–58.

22. Guzmán, M.R., Liquier, J., Brahmachari, S.K., Taillandier, E. (2006) Characterization of parallel and antiparallel G-tetraplex structures by vibrational spectroscopy. *Spectrochim. Acta. A. Mol. Biomol. Spectrosc.* **64**, 495–503.
23. Kodíček, M., Karpenko, V. (2013) *Biofysikální chemie* (Academie, Praha).
24. Ennifar, E. ed. (2015) *Nucleic Acid Crystallography: Methods and Protocols* (Humana Press).
25. Asherie, N. (2004) Protein crystallization and phase diagrams. *Methods* **34**, 266–72.
26. Wlodawer, A., Minor, W., Dauter, Z., Jaskolski, M. (2008) Protein crystallography for non-crystallographers, or how to get the best (but not more) from published macromolecular structures. *FEBS J.* **275**, 1–21.
27. Krauss, I., Merlino, A. An Overview of Biological Macromolecule Crystallization. *Int. J. Mol. Sci.* **2013**.
28. Cowtan, K. (2001) in *eLS* (John Wiley & Sons, Ltd).
29. Adrian, M., Heddi, B., Phan, A.T. (2012) NMR spectroscopy of G-quadruplexes. *Methods San Diego Calif* **57**, 11–24.
30. Kypr, J., Kejnovská, I., Renčíuk, D., Vorlíčková, M. Circular dichroism and conformational polymorphism of DNA. *Nucleic Acid Res.* **2007**.
31. Abriata, L. Structural database resources for biological macromolecules. *Brief. Bioinform.* **2017**, 659–669.
32. Singh, P.K., Bourque, G., Craig, N.L., Dubnau, J.T., Feschotte, C., Flasch, D.A., Gunderson, K.L., Malik, H.S., Moran, J.V., Peters, J.E., Slotkin, R.K., Levin, H.L. (2014) Mobile genetic elements and genome evolution 2014. *Mob. DNA* **5**, 26.
33. Kazazian, H.H.J., Moran, J.V. (2017) Mobile DNA in Health and Disease. *N. Engl. J. Med.* **377**, 361–70.
34. Chalopin, D., Naville, M., Plard, F., Galiana, D., Volff, J.-N. (2015) Comparative Analysis of Transposable Elements Highlights Mobilome Diversity and Evolution in Vertebrates. *Genome Biol. Evol.* **7**, 567–80.
35. International Human Genome Sequencing Consortium (2001) Initial sequencing and analysis of the human genome. *Nature* **409**, 860.
36. Wicker, T., Sabot, F., Hua-Van, A., Bennetzen, J.L., Capy, P., Chalhoub, B., Flavell, A., Leroy, P., Morgante, M., Panaud, O., Paux, E., SanMiguel, P., Schulman, A.H. (2007) A unified classification system for eukaryotic transposable elements. *Nat. Rev. Genet.* **8**, 973.
37. Hickman, A.B., Dyda, F. (2015) Mechanisms of DNA Transposition. *Microbiol. Spectr.* **3**, MDNA3-0034-2014.
38. Kapitonov, V.V., Jurka, J. (2008) A universal classification of eukaryotic transposable elements implemented in Repbase. *Nat. Rev. Genet.* **9**, 411–12.
39. Xiong, Y., Eickbush, T.H. (1990) Origin and evolution of retroelements based upon their reverse transcriptase sequences. *EMBO J.* **9**, 3353–62.
40. McClintock, B. (1950) The Origin and Behavior of Mutable Loci in Maize. *Proc. Natl. Acad. Sci. U. S. A.* **36**, 344–55.
41. Kim, N.-S. (2015) Transposable elements and genomics. *Genes Genomics* **37**, 111–12.
42. Feschotte, C., Pritham, E.J. (2007) DNA Transposons and the Evolution of Eukaryotic Genomes. *Annu. Rev. Genet.* **41**, 331–68.

43. Muñoz-López, M., García-Pérez, J.L. (2010) DNA Transposons: Nature and Applications in Genomics. *Curr. Genomics* **11**, 115–28.
44. Thomas, J., Pritham, E.J. (2015) Helitrons, the Eukaryotic Rolling-circle Transposable Elements. *Microbiol. Spectr.* **3**.
45. Siguier, P., Gournayre, E., Varani, A., Ton-Hoang, B., Chandler, M. (2015) Everyman's Guide to Bacterial Insertion Sequences. *Microbiol. Spectr.* **3**.
46. Vandecraen, J., Chandler, M., Aertsen, A., Houdt, R.V. (2017) The impact of insertion sequences on bacterial genome plasticity and adaptability. *Crit. Rev. Microbiol.* **43**, 709–30.
47. Charnavets, T., Nunvar, J., Necasova, I., Voelker, J., Breslauer, K.J., Schneider, B. (2015) Conformational Diversity of Single-Stranded DNA from Bacterial Repetitive Extragenic Palindromes: Implications for the DNA Recognition Elements of Transposases. *Biopolymers* **103**, 585–96.
48. Nunvar, J., Huckova, T., Licha, I. (2010) Identification and characterization of repetitive extragenic palindromes (REP)-associated tyrosine transposases: implications for REP evolution and dynamics in bacterial genomes. *Bmc Genomics* **11**, 44.
49. Nunvar, J., Licha, I., Schneider, B. (2013) Evolution of REP diversity: a comparative study. *Bmc Genomics* **14**, 385.
50. Di Nocera, P.P., De Gregorio, E., Rocco, F. (2013) GTAG- and CGTC-tagged palindromic DNA repeats in prokaryotes. *Bmc Genomics* **14**, 522.
51. Bertels, F., Rainey, P.B. (2011) Within-Genome Evolution of REPINs: a New Family of Miniature Mobile DNA in Bacteria. *Plos Genet.* **7**, e1002132.
52. Hickman, A.B., Dyda, F. (2016) DNA Transposition at Work. *Chem. Rev.* **116**, 12758–84.
53. Ivics, Z., Li, M.A., Mátés, L., Boeke, J.D., Bradley, A., Izsvák, Z. (2009) Transposon-mediated Genome Manipulations in Vertebrates. *Nat. Methods* **6**, 415–22.
54. van Opijnen, T., Camilli, A. (2013) Transposon insertion sequencing: a new tool for systems-level analysis of microorganisms. *Nat. Rev. Microbiol.* **11**, 435–42.

Svoluji k zapůjčení této práce pro studijní účely a prosím, aby byla řádně vedena evidence vypůjčovatelů.

Jméno a příjmení S adresou	Číslo OP	Datum vypůjčení	Poznámka